US Army Corps
of Engineers®

EM 200-1-7
1 February 2001

ENVIRONMENTAL QUALITY

PERFORMANCE EVALUATION (PE) PROGRAM

This manual is approved for public release; distribution is unlimited.

ENGINEER MANUAL

CEMP-RT/ CECW-E Engineer Manual 200-1-7	Department of the Army U.S. Army Corps of Engineers Washington, DC 20314-1000	EM 200-1-7 1 February 2001
	Environmental Quality PERFORMANCE EVALUATION (PE) PROGRAM	
	Distribution Restriction Statement Approved for public release; distribution is unlimited.	

AVAILABILITY

Electronic copies of this and other U.S. Army Corps of Engineers (USACE) publications are available on the Internet at http://www.usace.army.mil/inet/usace-docs/. This site is the only repository for all official USACE engineer regulations, circulars, manuals, and other documents originating from HQUSACE. Publications are provided in portable document format (PDF).

DEPARTMENT OF THE ARMY
U. S. Army Corps of Engineers
Washington, DC 20314-1000

EM 200-1-7

CEMP-RT
CECW-E

Manual
No.200-1-7

1 February 200 1

Environmental Quality
PERFORMANCE EVALUATION (PE) PROGRAM

1. Purpose. This Engineer Manual (EM) provides specific guidance, procedures, criteria, and tools for implementation of the U. S. Army Corps of Engineers (USACE) Performance Evaluation (PE) Program. Performance evaluation of analytical chemistry laboratories is recommended to ensure that technically competent and reliable laboratories are employed to generate analytical data of acceptable quality. This EM is intended for use by USACE personnel as a critical companion document to ER 1 1 10-1-263.

2. Applicability. This EM applies to HQUSACE elements, major subordinate commands, districts, laboratories, and field operating activities having responsibility for hazardous, toxic, and radioactive waste (HTRW) projects. This includes, but is not limited to, execution of the following programs: Defense Environmental Restoration Programs; Base Realignment and Closure; Superfund; Civil Works; Military Construction; installation environmental compliance; Defense Logistics Agency; Department of Energy; work for others; and any construction projects involving hazardous, toxic, and radioactive waste.

3. Distribution Statement. Approved for public release and unlimited distribution.

4. References. References are provided in Appendix A.

5. Discussion. This EM provides guidance for performance evaluation of analytical chemistry laboratories before contract awards and/or during sample analysis. The manual provides detailed guidance on design, preparation, certification, and use of PE samples needed to ensure fulfillment of the requirements of analytical chemistry aspects of the USACE Environmental Quality Assurance (QA) Program as prescribed in ER 1110-1-263 and ER 1180-1-6.

FOR THE COMMANDER:

ROBERT L. DAVIS
Colonel, Corps of Engineers
Chief of Staff

6 Appendices
(See Table of Contents)

DEPARTMENT OF THE ARMY
U. S. Army Corps of Engineers
Washington, DC 20314-1000

EM 200-1-7

CEMP-RT
CECW-E

Manual
No. 200-1-7

1 February 2001

Environmental Quality
PERFORMANCE EVALUATION (PE) PROGRAM

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1
Introduction

1.1. Purpose.

This manual provides details on the United States Army Corps of Engineers (USACE) procedures for the design, production, certification, and use of Performance Evaluation (PE) samples in the USACE Environmental Quality Assurance (QA) Program. PE samples are an important part of laboratory validation to ensure that quality data are obtained according to proper Chemical Data Quality Management (CDQM) activities. Quality of chemical data is critical to decision making for USACE environmental compliance and restoration programs. (See USACE 200-1-1, "Validation of Analytical Chemistry Laboratories," for details on lab validation and EM 200-1-6, "Chemical Quality Assurance for HTRW Projects," for CDQM information.)

Specifically, the PE samples discussed in this manual are used to:

- Test the ability of a laboratory to generate data of specified quality required by projects or regulators.
- Evaluate initial and continuous laboratory performance related to a contract/project.
- Ensure quality compliance in testing and maintaining overall data quality.
- Assess selection of appropriate analytical methods and laboratories.
- Assure that decisions affecting public health are based upon accurate and precise data generated by reliable laboratories.
- Meet qualifications for contractual purposes.
- Assist laboratories to improve their overall performance over time.
- Meet requirements of certification.
- Monitor laboratory performance on a routine basis.

1.2. Applicability.

This manual applies to Headquarters USACE (HQUSACE) elements, major subordinate commands (MSC), districts, laboratories, and field operating activities (FOA) having responsibility for in-house or contracted projects involving chemical testing of environmental media or chemical wastes. This includes, but is not limited to, execution of the following programs: Defense Environmental Restoration Program (DERP), Base Realignment and Closure (BRAC), Installation Environmental Compliance, Military Construction, Superfund, Civil Works, work for others including Defense Logistics Agency (DLA), Department of Energy (DOE), etc., and any other construction projects involving hazardous, toxic, and/or radioactive wastes.

Pertinent manual contents for various groups are listed below:

- PE Program Manager and PE Sample Suppliers: manufacture and use of PE samples.
- District customers, HQUSACE, and laboratories: how to request PE samples and participate in proficiency testing.
- Other branches of Department of Defense (DOD), other Federal agencies, and interested parties (commercial laboratories): scientific validity and legal defensibility of the PE samples and proficiency testing.

1.3. Distribution Statement.

This manual and any future revisions shall be approved for public release and unlimited distribution.

1.4. References.

References are listed in Appendix A.

1.5. Manual Overview and Scope.

This manual consists of six chapters and six appendices.

Chapter 1 presents an overview.

Chapter 2 introduces types of PE samples currently available from the USACE.

Chapter 3 discusses PE sample design.

Chapter 4 describes the manufacturing process of PE samples.

Chapter 5 describes PE sample certification process.

Chapter 6 describes the use of PE samples in the USACE Environmental QA Program.

The USACE Environmental PE Program is intended to cover all types of environmental analyses including chemistry, radiochemistry, biology, and microbiology in various environmental matrices. However, this manual applies primarily to chemistry proficiency testing of water and soil. The manual will be updated when PE samples for other types of environmental analyses are developed and used.

1.6. Acronyms and Definitions.

See Appendix B for acronyms, abbreviations, and symbols. Definitions comply with International Organization for Standardization (ISO), American National Standards Institute (ANSI), American Society for Quality (ASQ), and American Society for Testing and Materials (ASTM) usage.

The remainder of this chapter describes responsibility of groups in the PE program, fees and funding, and manual amendment.

1.7. Roles, Responsibilities, and Qualifications.

Distinct roles and responsibilities of five major parties in the PE Program are described below including Program Oversight Authority, Program Manager, PE Sample Providers, Customers, and Participating Laboratories.

1.7.1. Program Oversight Authority. The HQUSACE in Washington, DC, is responsible for oversight of the PE Program and final approval of major policies and operating procedures of the USACE Environmental Laboratory Validation Program. The HQUSACE annually determines the PE program budget and funding sources.

1.7.2. Program Manager. The Program Manager must have the technical expertise, administrative capacity, and financial resources to implement and operate a proficiency testing program. The qualifications of the Program Manager shall comply with the requirements of ISO 10011-2. The Program Manager shall follow the guidelines of ISO 10011 and 58 to inspect PE Sample Providers, at a minimum, on a biennial basis.

The USACE Hazardous, Toxic and Radioactive Waste Center of Expertise (USACE HTRW-CX) in Omaha, Nebraska, manages the USACE Environmental Laboratory Validation Program. Responsibilities include program development and implementation, PE Sample Provider approval, authorization for design and production of new and existing PE samples, and coordination of the PE Program participants.

1.7.3. Customers. USACE Project Managers or Contracting Officer Representatives (PM/COR) are the main customers of the PE Program. PM/COR from other Federal government agencies are potential customers. The customers are responsible for initiating laboratory validation processes and establishing project-specific Data Quality Objectives (DQO), which are considered when specifying PE samples. Customers may also be responsible for providing project-specific funding to support general laboratory validation activities including proficiency testing with PE samples.

1.7.4. PE Sample Providers. The quality system, testing facilities, and operating procedures of the PE Sample Providers shall conform with ISO 9000 series of standards and Guides 25 (17025), 34, 35, and 65. PE Sample Providers involve several government agencies including two environmental laboratories under USACE Environmental Research and Development Center (ERDC) and United States Environmental Protection Agency (USEPA). The PE Program may be supplemented by commercial sources. The Providers are responsible for the design, production, certification, and distribution of PE samples that meet the guidelines or requirements described in this manual and provide technical support to the Program Manager.

1.7.5. Participating Laboratories. Participating laboratories shall conform with ISO 9000 series of standards, ISO Guide 25 (17025), and USACE EM 200-1-1 requirements. Laboratories involved in proficiency testing shall explicitly follow the PE sample analysis requirements and data reporting schedule (see Section 6.1.5 of Chapter 6). Participating laboratories are responsible for PE sample analysis and data reporting costs.

1.8. Fees, Budget, and Funding.

1.8.1. Fees. The unit cost of PE samples includes material and labor costs for development, production, and shipment of PE samples; evaluation of PE results; and preparation of PE reports. A fee schedule is issued each fiscal year on PE unit costs that may be billable to individual customers.

1.8.2. Budget Request and Funding. Budget requests to HQUSACE should be made annually by the Program Manager and include proposed tasks such as PE program operation and PE sample development. Depending on environmental programs, funding is on a yearly or project-specific basis.

Mixed funding for proficiency testing will be used if multiple funding sources are available. For projects with unavailable funds from the Program Manager office or non-USACE entities, customers are responsible for costs except for PE sample analysis and data reporting (participating laboratory's responsibility).

1.9. Effective Date and Amendment.

This manual is effective upon approval by the HQUSACE and shall remain in effect until superseded or terminated.

The manual's guidelines are supported by standard operating procedures (SOPs) that reflect specific day-to-day scientific and business procedures. These SOPs are prepared by the participants of the USACE Laboratory Validation Program and are updated frequently.

Chapter 2
Characteristics of PE Samples

2.1. Definitions of PE Samples.

A PE sample is a physical sample that is representative of a relevant matrix and contains one or more analytes of interest at concentration levels unknown to laboratories being evaluated. PE samples are used to evaluate the analytical and reporting performance of a laboratory under prescribed conditions against a given set of criteria.

2.2. Purposes of PE Samples.

The USACE Environmental QA Program uses PE samples to evaluate laboratory performance before a contract award and to continually monitor laboratory performance after a contract award. PE samples are not used as the sole tool for determining laboratory performance on sample analysis or data reporting, but as an integral element of the USACE Environmental QA Program. PE samples serve not only as a distinct and effective indicator for laboratory performance and data quality, but also as an efficient and effective tool to assist laboratories in improving data quality over time.

2.3. Types of PE Samples.

2.3.1. Fields of Testing. There are several ways to classify PE samples according to analytical parameters, target analytes, sample matrices, analytical methods, preparation techniques, method of applications, regulatory programs, etc. The USACE's PE samples are classified according to field of testing which are organized as "matrix-method-parameter/analyte" combinations (e.g., volatile organic compounds in water by Methods 5030B/8260B, lead in soil by Method 6010B.)

2.3.1.1. Matrix. The matrix of PE samples is generally classified as water or soil based on sample preparation methods. Most PE samples sent by the USACE are prepared with fortified reagent water or clean natural matrices that do not contain contaminants at detectable levels. Well characterized real-world matrices that contain analytes of interest are also used to prepare PE samples.

2.3.1.2. Method. The method of the matrix-method-parameter/analyte combinations indicates the sample preparation and analytical methods to be evaluated. PE samples are generally developed for evaluation of one or more preparation and analytical methods.

2.3.1.3. Parameter/Analyte. The parameter/analyte is usually a project-specified, method-listed, or program-regulated analytical parameter, target analyte, or group of target analytes.

2.3.2. Blindness. Based on the degree of blindness of the identity of PE samples, PE samples are classified as single or double blind.

2.3.2.1. Single Blind. A single blind PE sample is identified as a PE sample, but its composition is unknown to the laboratory. Compared with double blind PE samples, single blind PE samples are easier, less expensive to prepare and use, and are the most commonly used by the USACE Environmental QA Program.

2.3.2.2. Double Blind. Both the identification and composition of a double blind PE sample are unknown to the receiving laboratory. Double blind PE samples mimic field samples in analytes, concentrations, matrices, interferences, packaging, and related sample documentation. Double blind PE samples are indistinguishable from routine field samples so that a laboratory will not bias analytical performance or reports.

Which sample type is preferred? In certain situations, double blind PE samples are the better choice for assessing laboratories. However, single blind PE samples are the more practical and economical choice for the following reasons:

- The instability of water PE samples and the heterogeneity of soil PE samples are major concerns in double PE sample production and use.
- It is costly and difficult to obtain statistically valid acceptance criteria for project-specific double blind PE samples.
- Making PE samples indistinguishable from routine samples is a substantial obstacle.

2.3.3. Packaging. Based on sample preparation and packaging, PE samples can also be classified as full-volume or ampule.

2.3.3.1. Full-Volume. Full-volume PE samples mimic real-world field samples in composition, amount, and packaging. Because they are handled like real-world field samples in terms of log-in, storage, preparation, analysis, and reporting, they provide an evaluation of the entire operation of a laboratory from sample receiving to data reporting. However, it is more difficult and expensive to prepare and use full-volume PE samples due to their short shelf-lives. The majority of USACE's PE samples are full-volume PE samples that are freshly prepared prior to shipment.

2.3.3.2. Ampule. Ampule PE samples are concentrates sealed in glass ampules. Laboratories follow special instructions to dilute the concentrates prior to sample preparation and analysis. Because the concentrates are flame sealed, the shelf-lives of ampule PE samples are usually much longer than those of full-volume PE samples. A large quantity of ampule PE samples is prepared and distributed for interlaboratory round-robin studies. The USACE uses ampule PE samples for a few less stable PE

samples such as volatile organic compounds in soil and organophosphorus pesticides in water. Ampule samples can only be used as single blind PE samples.

2.4. Requirements of PE Samples.

Guidelines for PE sample composition and certification are described below.

2.4.1. Sample Composition. Ideally, PE sample analytes and matrices should be designed and selected based on certain key aspects of a specific project such as project DQO, site contaminants, analytical methods required, etc. Because of variations in DQO or site contaminants, total site-specific PE samples may be unavailable in a timely or cost-effective manner. It may be more practical to develop generic PE samples that resemble various actual field samples and provide a wide spectrum of challenges to all laboratories.

2.4.2. Sample Certification. Because PE sample results can be used to disqualify a laboratory from being awarded a contract or to reject the analytical data produced by a laboratory, the quality of PE samples should be scientifically valid and legally defensible. To ensure the quality of samples, the following objectives must be met:

- The samples must be carefully designed and prepared with traceability to national or international reference standards and with proper documentation for legal defensibility.
- The acceptance limits of PE samples should be based on the results of interlaboratory studies and/or published method performance information.
- PE samples from different production batches or designs must offer a consistent challenge. PE samples must be of high quality, well documented, homogeneous, stable, and affordable.

Chapter 3
Design of PE Samples

This chapter outlines general and specific design considerations for PE samples including customizing samples.

3.1. General Design Considerations.

General design considerations for PE samples include the following:

- Determine the use of PE samples early in project planning to allow adequate time for selection or design of samples.
- Define clear goals for PE samples around the project's analysis needs and DQO.
- Design PE samples so that the entire laboratory is evaluated according to PE sample goals.
- Provide consistent but also project-specific challenges to all participating laboratories.

PE samples should have these general characteristics:

- Have physical, chemical, and behavioral similarities to field samples to provide an accurate test of the laboratory's procedures and method manipulations.
- Provide an evaluation of laboratory proficiency in sample analysis and quality control.
- Be homogeneous, reproducible, and stable over a required time period.
- Be appropriately represented in terms of defensible acceptance limits.
- Have scientifically valid and legally defensible certification data (for each PE sample).
- Be available on a long-term basis in sufficient and reliable supply.
- Have production methods that are not time consuming or costly.

3.2. Specific Design Considerations.

Project-specific PE samples are ideal; however, they may not be cost effective, timely to produce, or available. Therefore, generic PE samples that meet the majority of most project needs should also be considered. The remainder of this chapter discusses the following design or selection considerations:

- Matrices
- Methods
- Parameters/Analytes
- Concentrations
- Quality Assurance/Quality Control
- Homogeneity, stability, and reproducibility
- Amount

- Safety
- Cost
- Traceability, statistical design, and documentation

3.2.1. Matrices. Sample matrices may be categorized into air, water, soil, sediment, sludge, ash, oil, waste, tissue, etc. The PE samples used by USACE are mainly prepared with reagent water, clean soils, or real-world environmental matrices. Design guidelines for matrices include the following:

- Base matrix type on the sample preparation methods required for field samples and the problem at hand.
- Consider the origin, mineralogy, and pretreatment of field samples when PE samples are prepared with clean soil or real-world matrices because significant matrix differences can exist.
- Remember that USACE will also prepare PE samples with site-specific sample matrices, such as spiked, well characterized field matrices by request. However, site-specific PE samples are usually not cost effective and may not be available in a timely manner.

3.2.2. Methods. General method considerations include the following:

- Use only preparation and analytical methods that provide equivalent results. Not all sample preparation and analytical methods are equal in performance, so the acceptance limits for a PE sample based on one set of preparation and analytical methods may not be applicable to another set of preparation and analytical methods.
- Demonstrate the equivalency of method performance on each target analyte if other methods must be used. See USEPA guidance presented in USEPA/530/SW-87/003, "Test Method Equivalency Petitions," to evaluate the equivalency of test methods. Specific guidelines for analytical and preparation methods are described below.

3.2.2.1 Analytical Methods. Decide on the analytical method or instrumentation for analysis according to the nature of the field samples and sensitivity requirements. For example, analytical methods for metal PE samples depend on the metal elements, concentrations, and project requirements. They include:

- Inductively coupled plasma-mass spectrometry (ICP-MS).
- Inductively coupled plasma-atomic emission spectroscopy (ICP-AES).
- Flame atomic absorption spectroscopy (FLAA).
- Graphite furnace atomic absorption spectroscopy (GFAA).
- Cold vapor atomic absorption spectroscopy (CVAA).
- Gaseous hydride atomic absorption spectroscopy (GHAA).

PE samples prepared for a highly sensitive method, such as ICP-MS or GFAA may not be appropriate for a less sensitive method, such as ICP-AES or FLAA. Because most USACE environmental projects request USEPA SW-846 methods, the majority of USACE PE samples are designed and prepared for these methods. The same PE samples designed for SW-846 methodology can also be used for other methodologies prescribed by USEPA Superfund, Drinking Water, or Waste Water Programs. Also, a properly designed PE sample may be used for proficiency testing of multiple methods of different analytical techniques such as gas chromatography (GC) and gas chromatography/mass spectrometry (GC/MS) analyses.

3.2.2.2. Preparation Methods. For each matrix-method-parameter/analyte combination, there are a number of applicable sample preparation methods including sample digestion, extraction, cleanup, and concentration procedures. Factors for choosing a method include the following:

- Different sample preparation methods, such as sonication versus Soxhlet or hot plate versus microwave, may result in different recoveries of target analytes from certain sample matrices.
- PE samples designed for one specific sample preparation method may not be appropriate for performance evaluation of another sample preparation method.
- Low target analyte concentrations caused by lower recoveries with certain preparation methods may make the PE samples not challenging or too challenging.
- The recoveries of target analytes in PE samples, especially those prepared with real-world matrices, should be empirically checked to ensure that the same analyte recoveries are achieved with different sample preparation techniques. Determine an acceptance range to encompass the diversity of techniques. In case of substantial differences in analyte recoveries among preparation techniques, set separate acceptance ranges for individual or groups of techniques.

3.2.3. Parameters/Analytes. Guidelines for PE sample parameters/analytes include:

- Make sure that a single blind PE sample of a matrix-method-parameter/analyte combination contains all or the majority of representative, method- or project-specified target analytes. The minimum number of analytes that will be present in each aqueous PE sample shall comply with the National Environmental Laboratory Accreditation Conference (NELAC) standards. A double blind PE sample, however, should not contain all target analytes because this would be so unusual as to make it recognizable as a PE sample.
- Include components that cause known analytical and preparatory interferences besides target analytes. This approach will uncover whether a laboratory performs adequate interference corrections.
- Consider as much information as possible about analytes of interest, possible interferences, and limitations of analytical methods.

Specific information about parameters/analytes in the rest of this section includes challenging analytes, false positives and false negatives, formulation strategy, and incompatible analytes.

3.2.3.1. Challenging Analytes. It is common to include both more and less challenging target analytes in a PE sample and to evaluate the performance of a laboratory on an analyte-by-analyte basis. Challenging analytes are unstable, reactive, or interfering under certain sample preparation and analysis conditions. Use these analytes to check whether a laboratory takes proper precautions and corrective actions. Examples include: breakdowns of DDT and Endrin in dirty injection ports of a GC; loss of dichlorobenzene, the most volatile of the semivolatile compounds, because of poor nitrogen blow-down technique; low recovery of phenols because of incomplete acidification of the sample, less than required extraction time, excess drying of the extract, etc.

3.2.3.2. False Positives and False Negatives. False positives or negatives are serious but common problems with environmental chemical analysis. Laboratory contamination and retention time shift are common causes of false positives and negatives of organic analysis, respectively.

Design PE samples to test the ability of a laboratory to avoid reporting false positives or negatives in these ways:

- Identify false positive problems by looking for detection of analytes that are purposely left absent.
- Identify false negative problems by adding low-level analytes and watching for non-detects. Or the PE samples may contain isomers of analytes that elute closely and possess certain common mass ions, high levels of transition metals that exhibit potentially interfering spectral lines, or excess phthalate esters or elemental sulfur that interferes with the analysis of pesticides or polychlorinated biphenyls.

3.2.3.3. Formulation Strategy. Use specific modifications of PE sample compositions to provide additional checks of the entire analytical process. For examples, see the following list:

- Semivolatile PE samples should contain acid, base, and neutral analytes over the full retention time range.
- The addition of isomeric pairs to organic PE samples checks GC resolution.
- The addition of phthalates to pesticide PE samples tests extract cleanup methods.
- The addition of oil to soil PE samples verifies whether a gel permeation chromatographic cleanup was performed as contract required.
- The use of potassium ferricyanide, instead of potassium cyanide, to prepare aqueous cyanide PE samples checks whether distillation was conducted.
- Various other analytes may be added to gauge instrument performance, such as: addition of chloromethane to volatile PE samples to check if correct purge flow was used; addition of di-n-octyl phthalate to semivolatile PE samples to determine if the temperature of GC/MS transfer line

was set too low; use of specific xylene isomers to indicate if proper standards and response factors were used to set up instrument criteria; etc.

3.2.3.4. Incompatible Analytes. Certain groups of compounds should not be combined since they will react together. Select specific reagents for each analyte on the basis of not only quantity and availability but also on chemical characteristics (i.e., stability and reactivity). Use these guidelines:

- High concentrations of semivolatile acids and bases should not be combined because they will react with each other causing subsequent loss of analytes.
- Silver and low-to-medium levels of chloride ions are incompatible and should not be mixed.
- Certain compounds may not even be compatible with some instruments and should not be used. For example, it is difficult to use GFAA to analyze PE samples with high concentrations of chloride ions because of suppression of analyte signals.
- An expiration date must be established and specified for all prepared materials including PE samples, reagents, reference standards, etc.
- PE samples that contain unstable target analytes may need to be supplied in sealed ampules or as a concentrate in nonaqueous solvents to enhance the stability of the analytes.

3.2.4. Concentrations. Concentration of PE samples should be determined using these guidelines:

- Concentrations should be near the levels expected in field samples or span the range of the analytical method.
- Concentrations may also be prepared near project or regulatory action levels to check laboratories' performance near action levels.
- Concentrations also depend on the target analytes of interest and the method of analysis. For example, the concentrations of metal PE samples depend on the metal elements, analytical methods, and project requirements. *As a general rule, to check a laboratory's performance in accuracy, the concentrations of target analytes should be at least ten times higher than sample-specific Method Detection Limits (MDL) to avoid excessive random errors near the MDL.* However, to check a laboratory's ability to avoid false negatives, the concentrations of target analytes may be set at three to five times above the MDL.
- Both high and low concentrations may be included to check if a laboratory would analyze field samples at project-required dilutions. Such multiple dilutions or concentrations may be needed to quantitate all target analytes correctly. However, avoid excess interferences between closely eluted compounds of substantial concentration difference that would require unique or non-routine treatments to pass PE samples.

3.2.5. Quality Assurance/Quality Control. Use proper design and different types of PE samples to detect and correct specific quality assurance/quality control (QA/QC) problems.

- Analytical precision can be verified by duplicate analyses of well homogenized PE samples (e.g., aqueous PE samples that are intrinsically homogeneous down to a molecular level) that contain many analytes at midrange concentrations.
- Matrix spike recovery problems can be verified by submitting a spiked field sample and a spiked extract or digest of the same field sample. Differences in recoveries between pre- and post-extraction/digestion spikes will demonstrate whether the extraction/digestion process at a laboratory is performing well.
- Precision data based on PE samples of clean matrices and real-world matrices provide information about the true laboratory precision against the precision difficulty associated with the method on complex matrices.

3.2.6. Homogeneity, Stability, and Reproducibility. In general, USACE's PE samples are individually prepared for each laboratory on an as needed basis. To assure homogeneity in representative subsamples, follow ASTM Standards D6051 and D6323. Sample homogeneity depends on the following:

- Specificity of the characteristic.
- Precision of measurement. Because aqueous PE samples are prepared by spiking target analytes into individual sample containers, reproducibility of PE samples within and between preparation batches is critical. The heterogeneity of aqueous PE samples is mostly due to differential contamination during the preparation and final packaging or incomplete dissolution or equilibration of an analyte(s) in the spiking solution.
- Defined sample size of the test portion. PE samples of solid matrices are often prepared in large quantity and subsampled for individual PE samples. Homogeneity and stability of the bulk PE materials are crucial to ensuring that equivalent samples are sent to all participating laboratories over an extended time period. Soil PE samples are heterogeneous in composition by nature but are accepted as homogeneous with respect to a specific characteristic if no difference on this characteristic among different parts can be experimentally detected.

The remainder of this section describes initial verification and ensuring equivalent PE samples for aqueous, solid, and unstable sample types.

3.2.6.1. Initial Verification. PE samples must be verified on a regular basis to show that there are no significant differences on the prepared analyte concentrations within and between batches. Prior to PE sample use follow these guidelines:

- Use generally accepted testing procedures to establish the homogeneity, stability, and reproducibility of PE samples. A suitable testing procedure should compare PE samples across production runs for comparability at $\alpha = 0.05$ level using appropriate statistical testing.

- Refer to ISO Guide 35 for Certification of Reference Materials and ASTM Standards E826, D4515, D4841, and F1469 for suitable testing procedures.

3.2.6.2. Verification of Aqueous PE samples. Full volume, aqueous PE samples are freshly prepared on the day of shipment. The samples are considered stable within the method specified holding time under method-recommended storage conditions. Test guidelines for verifying aqueous PE samples include:

- Use a single holding time study for each particular type of aqueous PE sample according to ASTM Standards D4515 and D4841.
- Verify PE samples that contain unstable analytes or analytes at low concentrations with "end of holding time" sample analyses. The sample is considered stable within the holding time if the mean measured concentrations of "end of holding time" samples do not show significant differences (i.e., at $\bullet = 0.05$ using a conventional *t* test) from the mean prepared concentrations based on pre-shipment verification analysis.

Ideally, PE samples should be stable for some time after the return of results by participating laboratories in case there are any queries and problems that need to be addressed.

3.2.6.3. Verification of Solid PE Samples. For soil and sediment PE samples, following these guidelines for monitoring stability:

- Verify the sample concentrations of bulk PE sample materials on a routine basis; results should fall within the 95% confidence intervals (i.e., $\pm 2\bullet$) of previously determined mean reported values.
- Carry out stability studies (depending on the degradation mechanisms) at elevated temperatures to accelerate the degradation rate and hence reduce the time needed to collect sufficient data.
- Obtain stability data from the manufacturers of materials or through previously published data.
- Refer to the detailed procedures for investigations of sample homogeneity, stability, and reproducibility in Section 5.3 of Chapter 5.

3.2.6.4. Unstable PE Samples. The following steps are taken for unstable PE samples to ensure equivalent PE samples:

- Provide special instructions on sample storage and treatment to participating laboratories.
- Take precautions to ensure that samples remain unchanged, at a minimum, till the end of method-specified holding time.
- Package PE samples that consist of a mixture of powders of different relative density or of different grain size carefully to avoid segregation that may result during transport.
- Seal PE samples into ampules if reactions with the atmosphere may be expected. The ampule may be filled with an inert gas.

3.2.7. Amount. Use these guidelines to determine PE sample amount:

- Provide a limited amount of PE samples so that a laboratory has only one chance to prepare and analyze them and cannot split samples and compare results.
- Use "full-volume" PE samples (unless stability is a concern) that mimic the compositions of field samples to evaluate the entire laboratory (sample handing, preparation, analysis, method selection, record keeping, and data validation and reporting).
- Use ampule PE samples of reduced volume or mass but at high concentrations if stability is a major concern. For example, use ampule PE samples for organophosphate pesticides in water and volatile organic compounds in soil.

3.2.8. Safety. Because hazardous and toxic chemicals are used to prepare PE samples, safety guidelines to consider include:

- Establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.
- Consider any hazards that the PE samples might pose and communicate them to parties that might be at risk.
- Make Material Safety Data Sheets (MSDS) available for each material and read and follow them carefully.
- Treat all unknown samples as hazardous and toxic, and handle all toxic materials in a well ventilated fume hood.
- Use a Chemical Hygiene Plan (CHP) to establish responsibilities, policies, and procedures for handling hazardous chemicals in the laboratories. (Both PE Sample Providers and participating laboratories should have a CHP.) The CHP shall comply with Occupational Safety and Health Administration (OSHA) Standard, 29 Code of Federal Regulations (CFR) 1910.1450. This EM does not purport to address all of the safety concerns, if any, associated with its use.

3.2.9. Cost. Proficiency testing is an effective but also expensive way to evaluate laboratory performance. The cost includes, but is not limited to, operation, maintenance, instrument downtime and repair costs, etc. Ways to reduce the cost of PE sample preparation include:

- Consider the potential use of a single PE sample for evaluation of multiple methods. For example, a well designed PE sample for volatile organic compounds could be used to test laboratory performance on multiple analytical methods including Methods 601, 602, 8021B, and 8260B.
- Select analyte, concentration, and packaging so that one PE sample can serve multiple purposes successfully at substantial cost reductions. For example, with proper packaging, a single PE sample may be used as a single blind or double blind PE sample.

3.2.10. Other Considerations. Other design guidelines discussed in this section include traceability, statistical design and analysis, and documentation and record keeping.

3.2.10.1. Traceability. The assigned purity or concentration value and the associated uncertainty of all chemicals, reagents, and reference standards used for preparation of PE samples should be traceable to National Institute of Standards and Technology (NIST) or other equivalent standards. The following guidelines are necessary for PE Sample Providers to follow:

- Provide evidence of the traceability for all reagents, chemicals, and standards used for PE sample preparation to known and accepted national or international certified reference materials.
- Have all measurement and testing instruments, equipment, and apparatus certified or calibrated against national or international standards to ensure the accuracy of measurements.

3.2.10.2. Statistical Design and Analysis. Statistical design and analysis play a key role in a PE Sample Program. Involve a statistician or someone with extensive statistical design and analysis experience in experiment design and data analysis of PE samples. Expected responsibilities of the statistician include:

- Participate early in the planning, design, and certification of PE samples.
- Perform critical examination of experimental data to identify and treat data gaps, outliers, or other irregularities.
- Test the suitability of the model.
- Analyze data with appropriate statistical tools.

3.2.10.3. Documentation and Record Keeping. The PE Sample Providers must:

- Establish and maintain a secure record keeping system with limited access.
- Maintain a system for logging files in and out to suit any particular circumstances that might be needed to comply with applicable requirements.
- Retain all individual measurement observations, appropriate calculations and derived data, sample preparation, instrument calibration records, and certification procedures for a minimum of five years after the discontinuance of the PE samples. (Records may need to be retained longer per standards from ISO, ASTM, ANSI, ASQ, USEPA, and USACE.)
- Prepare PE sample documentation that is complete, accurate, legible, indelible, unambiguous, and objective such that any information necessary for interpretation and reconstruction of the PE samples is available.
- Maintain an original copy of all data reports showing all corrections or changes on file.

The trend to replace hard copy documentation with electronic documentation is expanding rapidly, and the data management practices used to protect the integrity of electronic data are becoming increasingly important. Minimum records to be maintained include:

- A description of the hardware and software used.
- Written standard operating procedures that document procedures for generating, reviewing, and validating electronic data.
- Results of periodic in-house audits of electronic data generation and reporting.
- Backup files for electronic data generated in accordance with USEPA Directive 2185, "Good Automated Laboratory Practices." This document suggests appropriate frequencies for generating backup files (including system operating files and application files) and suggestions for off-site storage of software and hardware backups.
- Written certification (for electronic certification with a unique identifier) by the laboratory manager verifying the authenticity of each data report. Maintain such certifications for audits.

3.3. Customization.

Upon request, the Program Manager will provide customized PE samples that are totally site-specific in sample compositions and double blind to the laboratory.

3.3.1. Site Specificity. An ideal PE sample should be site-specific; however, total site-specific PE samples may not be cost-effective and/or available in a timely manner. The majority of USACE aqueous PE samples are tailored to meet site specificity in target analytes and concentrations; however, most solid PE samples are not site-specific due to technical and cost considerations.

3.3.2. Double Blindness. To conceal the identity of double blind PE samples, follow these guidelines:

- Pay special attention to the constituents of the PE samples during the design.
- Make the constituents, including target analytes, concentrations, interferences, matrices, etc., compatible with field samples. If the constituents of PE samples are substantially different from those of field samples, laboratories could easily identify them as PE samples.
- Use the same packaging, shipping, and documentation procedures as those of the field samples to hide the PE sample identity.
- Insert double blind PE samples with the same packaging, labels, and documentation as those of the field samples into normal field sample streams.
- Use a fictitious contract to ship double blind PE samples.

Chapter 4
Preparation of PE Samples

Manufacturing PE samples of authenticated compositions requires careful planning so that PE samples achieve their intended purposes. The preparation procedures for USACE PE samples described in this chapter include the following:

- Matrix selection.
- Selection of chemicals and reagents.
- Facility and equipment requirements.
- General and specific preparation procedures including fortified, real-world, and commercial PE samples.
- Handling and storage requirements.
- Packing and shipping guidelines.

4.1. Selection of PE Sample Matrices.

Three types of sample matrices commonly used to prepare PE samples are: reagent water, clean soils, and real-world environmental matrices. Guidelines for matrix selection include the following:

- Use the sample preparation methods required for project samples to determine the matrix type of PE samples.
- Use water PE samples for laboratories involved in analysis of aqueous samples.
- Use both water and soil PE samples for laboratories involved in analysis of soil samples. It is USACE policy that a laboratory must pass both water and soil PE samples to be considered for validation of soil sample analysis.

4.1.1. Reagent Water Matrix. Use reagent water that is free of contaminants of concern at MDL for preparation of aqueous PE samples by following these guidelines:

- Prepare by distillation or other equivalent processes, then polish with a mixed bed of ion exchange materials, a 0.2-• m membrane filter, and/or an organic adsorption system.
- Use high-purity reagent water (ASTM Type I or II grade of reagent water) immediately after production; any kind of storage will cause some form of purity degradation.
- Monitor and document reagent water quality daily.

4.1.2. Clean Soil Matrix. Various types of soils that are free of contaminants of concern at MDL can serve as clean soil matrices. Select suitable clean soils according to target analytes, DQO, sample preparation, and analytical methods. In addition, some specific considerations include:

- Soil composition and texture impact preparation and PE sample performance. Coarse texture and large grain size affect homogeneity and representativeness of subsamples.
- High alkalinity affects the stability of acidic target analytes.
- Organic and metal contents may vary; however, avoid soils that are very reactive to acids or other reagents used for sample preparation.
- Contaminated soils may be used after purification with acid wash, solvent extraction, or thermal treatment. However, avoid excessive thermal treatment because it may destroy organic matter and soil structure. The resulting matrix would not be site-specific or similar to untreated soil. Refer to Appendix C for two procedures for soil cleaning.
- Background soils that contain minute amounts of indigenous contaminants and/or interferents may be more practical for PE sample preparation.
- Background soils collected from project sites mimic field samples.

4.1.3. Real-World Matrices. To mimic field samples, real-world environmental matrices, such as tap water, surface water, ground water, soils, and sediments can also be used to prepare PE samples. Use a well characterized, contaminated environmental matrix in one of three ways: as it is, fortify it with target analytes, or dilute it with a clean environmental matrix to desired analyte concentrations. The target analytes of real-world contaminated soils usually possess special speciation or partition in sample matrices or penetrate into the micropores of soil particles where they are tightly bound. These factors make the PE samples more challenging.

4.2. Selection of Chemicals and Reagents.

Considerations for chemical and reagent selection discussed in this section include purity, traceability, concentration, certified reference material, chemical composition, and matrix selection.

4.2.1. Purity. All chemicals shall be proven free of contaminants and interferences or be of sufficient purity to meet the intended purpose. Guidelines for purity include the following:

- Use reagents that conform to the specifications of the Committee on Analytical Reagents of the American Chemical Society (ACS). Other grades may be used if the reagent's purity permits its use without adverse impacts on the bias and precision of subsequent analyses.
- Use acids of reagent grade or better except for those used for ICP and GFAA work, which need to be of high-purity grade or equivalent.
- Use solvents of chromatographic grade or better.
- Document all critical information about chemicals and reagents.

4.2.2. Traceability. The quality and concentrations of all chemicals, reagents, and materials used for production of PE samples should be traceable to national or international standards. Document traceability of materials in these ways:

- Use unique identifiers for all purchased and in-house prepared reagents to establish and maintain traceability throughout PE sample manufacturing.
- Maintain a log book for all chemicals, reagents, and materials used for production.

4.2.3. Concentration. Determine the concentration or purity of all starting materials by various procedures that are listed below in an order of increasing uncertainty.

- True or known concentrations as determined by specific formulations (e.g., manufacture or dilution.)
- Certified reference values as determined by multiple definitive methods of different quantitation principles.
- Reference values as determined by analysis along with standards traceable to national reference standards.
- Consensus values as determined by a group of expert laboratories that have demonstrable competence in the analytical methods to be evaluated.

Use multiple definitive methods as defined below to minimize systematic errors in determining analyte concentrations. Definitive measurement methods are methods that have been rigorously proven and demonstrated to be accurate within specified limits. Definitive methods provide certified reference values that are directly traceable to fundamental units of measurements (e.g., mass, length, time, volt, ohm, etc.) or indirectly related to the fundamental units through physical or chemical theory expressed in exact mathematical equations. Theoretically, any methods of high precision and accuracy might be qualified as definitive methods for any combinations of analyte and matrix. In practice, however, a definitive method must have, at least, sufficient documentation to prove its accuracy and use indirect means to establish measurement errors in real sample matrices.

4.2.4. Certified Reference Materials. When commercially available reference materials are utilized as PE samples or used to prepare PE samples, follow these guidelines:

- Be sure that the acceptance limits for reference materials and methods being evaluated were established using the same or equivalent sample preparation and analysis methods. This is especially critical for commercially available solid reference materials where the certified concentrations are often the total concentrations instead of the recoverable concentrations given by most USEPA methods.
- Establish acceptance limits for target analytes in soil PE samples based on the same preparation and analysis methods to be evaluated. Otherwise, results may not be equivalent to those of the methods being evaluated.

4.2.5. Chemical Compositions. Determine chemical compositions using these guidelines.

- Consider the types of cation/anion pairs, the oxidation states of cations, the ligands, the hydrations, etc., for inorganic standards.
- Consult literature for solubility, stability, and chemical and physical properties of a standard, especially if the standard contains compounds that are not previously used in a chosen solvent.
- Collect current results of Karl Fisher analysis to determine the concentration of PE samples if the standard is hygroscopic.

4.2.6. Selection of Matrix. After matrix selection:

- Verify that a current purity assay of the matrix is on file. If one is not, determine and document the purity of the matrix with a minimum of two definitive methods of different measurement methodologies by a minimum of two reliable laboratories.
- Use reagent water meeting ASTM Type I for inorganic analyses and water meeting Type II for organic analyses.

4.3. Facility and Equipment Requirements.

Laboratories providing PE samples must meet these standards:

- USACE validation according to the requirements in EM 200-1-1.
- Compliance with ISO Guides 25 (17025) and 34, ISO 9000 series of standards, and ASTM Standard

D5522 for facilities, equipment, and quality systems used in PE sample preparation and testing.

Standard for facilities, balances, ovens, thermometers, pH meters, and volumetric glassware, are outlined below.

4.3.1. Facility. Facility requirements include the following:

- Provide laboratory security by limiting access and using a secured storage area.
- Monitor the laboratory for proper air flow, ventilation, humidity, temperature, and lighting.
- Provide adequate work space such that PE samples can be prepared and tested safely and efficiently in order to minimize the possibility of cross contamination.
- Provide a stable power supply, sufficient exhaust hoods, and proper storage facilities.
- Use adequate procedures and facilities for the collection, storage, and disposal of chemical wastes.

4.3.2. Balances. Standards for balances include:

- Use appropriate balance ranges for applications.
- Check each balance daily (or with use) with two ASTM Class 1 or equivalent weights that bracket the expected weight range. Check analytical balances monthly with a series of Class 1 weights and document the results. Analytical balances should have readabilities of 0.01 to 0.1 mg to meet measurement requirements.
- Correct any variance of greater than 0.1% between expected weights and actual weights.
- Check top-loading balances with, at a minimum, ASTM Class 2 or equivalent weights within ±1%. Top-loading balances should have readability of 0.001 to 0.01 g.
- Use Class 1 and 2 weights calibrated within the last five years and traceable to NIST reference standards.

4.3.3. Ovens and Furnaces. Guidelines for oven and furnace use include:

- Check the temperature of each drying oven before and after each use to verify the correct operating temperature for the given operation.
- Check and document the extent of temperature swings at different operating temperature ranges.
- Verify the temperature at least annually in the operating ranges for muffle furnaces.

4.3.4. Thermometers. Standards for thermometers include:

- Check the calibration of each mercury or alcohol thermometer at least annually against an NIST traceable thermometer. Check at two separate temperatures bracketing the expected temperature range.
- Calibrate electronic and dial-type thermometers at least quarterly against an NIST traceable thermometer.
- Use thermometers with ±1°C accuracy.

4.3.5. pH Meters. Guidelines for pH meter use include:

- Calibrate pH meters with buffer solutions prepared with NIST primary buffer salts or commercial secondary buffer solutions traceable to NIST's standards.
- Standardize the pH meter at least daily with two buffers that bracket the expected pH range. The bracket must be no more than three to four pH units wide.
- Use pH meters with ±0.1 pH unit accuracy.

4.3.6. Volumetric Glassware. Standards for volumetric glassware use include:

- Make all volumetric measurements with ASTM Class A glassware. ASTM Class A pipets and calibrated microsyringes must be used for delivering and spiking during preparation of PE samples. Variable pipetters can be used if they are gravimetrically verified to be in calibration; however, glass pipets are preferred.
- Use volumetric glassware that yields ±1% accuracy.

4.4. Preparation of PE Samples.

The goal of preparation is that PE samples resemble routine field samples in appearance, characteristics, analyte content, and concentration level. PE samples can be prepared in two ways:

- Spiking known amounts of analytes into a clearly defined homogeneous matrix (known as fortified PE samples).
- Defining homogenized real-world samples (known as real-world PE samples).

Both of these methods as well as commercially prepared PE samples are discussed in this section.

4.4.1. General Preparation Procedure. Regardless of the type of PE samples, the general procedure for preparing PE samples is outlined below. Some of these steps, such as (6), (7), (11), etc., may be exempted when additional batches of the same or similar PE samples are prepared.

(1) Determine matrix type, analytical method, and analytical instrumentation.
(2) Calculate the total amount of PE samples needed by volume or weight.
(3) Select analytes, interferences, solvents, and preservatives.
(4) Decide the concentration of each component.
(5) Select stock materials and calculate appropriate amounts to add.
(6) Write step-by-step instructions (i.e., SOPs).
(7) Perform an error analysis and define performance requirements.
(8) Obtain stock materials.
(9) Prepare PE samples.
(10) Verify the concentration of each component in PE samples.
(11) Verify the composition of PE samples by multilaboratory referee analyses.
(12) Establish performance acceptance limits of each PE sample.
(13) Characterize any indigenous level of target analytes and/or interferents when using real-world materials. Additional interlaboratory analyses are needed to establish the acceptance limits of real-world PE samples.

4.4.2. Fortified PE Samples. General guidelines regardless of matrix are given below followed by fortification techniques for aqueous and soil PE samples.

4.4.2.1. General guidelines. Prepare fortified PE samples of any matrix using these guidelines:

- Minimize variations between PE samples by preparing a large amount of PE sample and subsampling it to create individual PE samples.
- Preparing individual samples if analyte loss to containers is a major concern. For example, hydrophobic petroleum hydrocarbons tend to separate from water matrix and stick to the container wall. Therefore, the container must be rinsed with extraction solvent to enhance the recovery of petroleum hydrocarbons. So use individually spiked PE samples.
- Use experienced senior chemists to prepare spiked samples and improve reproducibility.
- Prepare aqueous PE samples of one liter for semivolatile organic analyses and 40 mL volatile organic analyses.
- Prepare aqueous PE samples for inorganic analyses from 200 mL to one liter, depending upon the target analytes and analytical methods.
- Prepare a trip blank to accompany volatile organic samples for each analytical method.

4.4.2.2. Aqueous PE Samples. Aqueous PE samples can be prepared by either gravimetric or volumetric methods depending on the physical properties of bulk materials.

- Use gravimetric methods for very volatile materials or those with unreliable density values.
- Use volumetric methods when bulk materials have low vapor pressures and reliable density values.
- Prepare aqueous PE samples on the day of shipment, usually early in the week to allow adequate preparation time for participating laboratories to perform digestions, extractions, cleanups, etc., before the weekend.

4.4.2.2.1. Gravimetric Method. A gravimetric method can be used for both liquid and solid fortifications. Follow these procedures for a gravimetric method:

- Determine the density of the solution so that volume-to-weight conversions can be calculated if weights will be used for calculation.
- Calculate the mass of starting materials to be weighed for the specified volume of solution. The prepared concentration should be within ±1% of the target concentration of the solution.
- Use this formula:

$$M\,(g) = \frac{V_s\,(mL) \times C\,(\mu g/mL)}{10^6\,(\mu g/g) \times P}$$

where: M = mass of the starting material required
V_s = volume of the final solution desired
C = concentration of the final solution desired
P = purity as a fraction of the starting material

Purity used in the calculation could be a certified value from the manufacturer or vendor.

- Perform weighing in a glove box flushed with high purity nitrogen gas to reduce the risk of contamination.
- Check the balance with NIST traceable calibration masses on the day the standard is prepared to ensure accuracy of weighing.
- Record the results in the balance logs.
- Guard against static charges on the balance or the beaker influencing the balance readings with an antistatic device.
- Remove the beaker from the balance and pipet some solvent (~3 mL) into the beaker to aid the dissolution and prevent evaporation of the weighed samples as soon as balance door is closed and weight is measured.
- Cover the top of the beaker to avoid contamination.
- Record the lot numbers of all reagents and starting materials.
- Designate a unique identification number for the solution.
- Date and initial the entry.

4.4.2.2.2. Volumetric Method. A volumetric method can only be used for liquid fortification. Follow these procedures for a volumetric method:

- Calculate the volume of each bulk required.
- Allow all solutions to warm to ambient temperature and mix them thoroughly to ensure homogeneity and accurate volumetric measurement. The prepared concentration should be within ±1% of the target concentration of the solution.
- Use this formula:

$$V\ (mL) = \frac{V_S\ (mL) \times C\ (\mu g/mL)}{d\ (g/mL) \times 10^6\ (\mu g/g) \times P}$$

where: V = volume of the bulk required
V_s = volume of the final solution desired
C = concentration of the final solution desired
d = density of the bulk
P = purity as a fraction of the bulk

- Obtain starting materials and stock solutions at appropriate concentration levels to minimize the amounts required and to remain in the realm of accurate measurements of laboratory ware. Avoid dilutions requiring odd sizes of volumetric ware.
- Keep volume measurements between 50 • L and 1 L and weights between 10 mg and 1 kg.
- Choose the more concentrated solutions of stock solutions to minimize potential contamination.
- Use minimal headspace and exposure to the atmosphere while preparing volatile liquids.
- Use diluents that already contain any required preservatives so that final volumes are not altered by preservation.

4.4.2.3. Soil PE Samples. Soil PE samples are prepared with clean soils in bulk or in small quantity. Use these guidelines for determining sample size.

- Use bulk quantity PE samples for stable target analytes and subsample for use over an extended time period.
- Use small quantity PE samples for less stable target analytes and prepare as individual PE samples on an as needed basis. Currently, small quantity PE samples are only used for explosives PE samples where preweighted, clean soils in extraction vials are spiked with target analytes.

4.4.2.3.1. Preparation of Clean Soils. To prepare clean soils, follow these steps:

- Collect a large quantity of soil from a clean site.
- Check soil cleanliness to ensure that no contaminants or interferents exist at concentrations above the detection limits of the analytical methods being evaluated.
- Remove rocks, sticks, and other foreign material and dry soil at ambient temperature to approximately 2 - 4% moisture.
- Grind, mix, and run dried soils through a fine screen to remove large particles.
- Determine grinding and mixing times by checking particle size and soil consistency after short interval runs.
- Sieved soil should be homogeneous with the consistency of fine powder and specific particle sizes.
- Use the sample mass required for PE samples and the maximum sampling errors allowed (15%) to determine maximum particle size (according to Pierre Gy's Sampling Theory). See Appendix D for the calculation of maximum allowable particle size that can be accommodated by a given sample mass at various sampling errors. (A No. 50 mesh sieve generally yields materials of adequate particle size.)
- Keep the overall sampling error as low as practically feasible but it should not exceed 15%.

4.4.2.3.2. Verification of Clean Soils. Use the following procedures to confirm the cleanliness of the prepared clean soil after processing.

- Randomly select multiple sets of clean soil samples for analysis by several reliable laboratories using the analytical methods being evaluated. An alternative is to have one reliable laboratory use several different analytical methods with similar or lower detection limits than those of reference methods.
- Instruct laboratories to prepare samples with different preparation methods known for good analyte recovery (within 70% to 130%).
- Accept the clean soil, if the concentrations of any native target analytes or interferents are below the MDL of the analytical methods to be evaluated or are negligible (i.e., <5%) in comparison with the concentrations of spiked target analytes. Otherwise, remove the trace contaminants if necessary with acid washing or solvent extraction, or find another, cleaner soil.
- Bring the pH value of any cleaned soil to neutral. Otherwise, certain spiked target analytes might be decomposed, transformed, and lost in recovery.

4.4.2.3.3. Fortification of Clean Soils. Clean soils can be fortified with target analytes by solid, liquid, or vapor fortification methods. The choice of a specific method depends on the chemical properties of target analytes and the stability of the PE samples.

4.4.2.3.3.1. Solid Fortification. Use solid fortification for analytes of low vapor pressure and solubility following these guidelines:

- Reduce the particle size of solid reagents to small particles of approximately the same size (ca. • 300 microns) before mixing. This makes blending easier, and components will be less prone to segregate during storage and transit.
- Use similar amounts of each component because it is very difficult to evenly distribute otherwise. If extremes in relative amounts cannot be avoided, do the blending in stages by spiking and blending a small quantity of the main component, then mixing and blending with the rest.
- Note that solid fortification works best if the specific gravities of the analyte compounds are similar to that of the host matrix (ca. 2.5 g/cm^3 for soil) within reasonable limits.

4.4.2.3.3.2. Liquid Fortification. Use liquid fortification for soluble and less volatile analytes following these guidelines:

- Spray analyte solutions over homogenized soil in small increments.
- Mix the soil thoroughly and spray again after vaporization of the solvent.
- Repeat until all analyte solutions are used up.
- Rinse the spray bottles with more solvent and spray over the soil again to ensure that all target analytes are quantitatively transferred to the homogenized soil.
- Mix thoroughly after drying to avoid analyte deposits.
- Improve mixing using enough solvent to form a runny paste or mud. Stir the paste occasionally while drying and, when completely dry, re-grind and blend.

- An alternative to the above method is to soak soils with a large volume of analyte solution, remove the solvent, and dry, grind, and blend the soil.

4.4.2.3.3.3. Vapor Fortification. Use vapor fortification for volatile organic analytes following the vapor fortification technique developed by Alan Hewitt of USACE Cold Regions Research and Engineering Laboratory (CRREL) in Hanover, New Hampshire. Steps in this technique include:

- Place ampules of preweighed, homogenized soils in a sealed desiccator with desiccant for a minimum of three days.
- Transfer soils to another desiccator containing fortification solution of target analytes.
- Allow the soils equilibrate with the vapor of analytes for a minimum of 14 days.
- Remove the ampules and flame seal them quickly.
- Use vapor-fortified PE samples by breaking open the ampule in a purge cell containing distilled water, appropriate surrogates, and internal standards.

4.4.3. Real-World PE Samples. Guidelines for the collection, preparation, and fortification of real-world PE samples are outlined in this section.

4.4.3.1. Collection. Real-world environmental materials that have various contaminants of concern at significant concentrations can be prepared, certified, and used as PE samples. Due to stability concern, only soils or sediments are usually used for preparation of real-world PE samples. Collect materials using these procedures:

- Locate candidate contaminated sites through solicitation or project documents.
- Collect samples and analyze to verify the location and extent of contamination prior to collection of a large quantity of materials.
- Avoid environmental materials that contain numerous analytes at low concentrations because they may cause problems with the assessment of laboratory performance.

4.4.3.2. Preparation. The preparation process for real-world PE samples is similar to that of clean soil PE samples. Procedures are listed below:

- Remove extraneous materials such as rocks, sticks, etc.
- Dry, grind, and mix materials with mills or grinders to a powdery consistency with fine particle sizes. The maximum allowable particle size depends on the maximum allowable subsampling error and subsample size (See Appendix D). In general, soil materials sieved through No. 50 mesh (• 300 microns) have adequate particle sizes for most environmental analyses including mercury analysis, which requires a sample size of 0.2 grams.

4.4.3.3. Fortification. The content of real-world PE samples can be modified by fortification with target analytes or by dilution with clean matrices. The fortification techniques previously described may be used. (See Section 4.4.2.3.3.) After fortification or dilution, additional grinding and blending steps are necessary to ensure the homogeneity of the final products.

4.4.4. Commercially Available PE Samples. Commercially available PE samples may also be used for proficiency testing. Guidelines for use include:

- Use commercial PE samples after evaluation and approval by the Program Manager. This evaluation is based on national or international standards and requirements issued by ISO, ASTM, and USEPA. The evaluation covers manufacturers' preparation and certification of PE samples.
- Inform laboratories that they must explicitly follow manufacturers' instructions.
- Use only certified reference materials (CRM). NIST is the most widely used supplier of CRM. However, using NIST values for solid materials may sometimes lead to comparison errors on data obtained with USEPA organic and inorganic sample preparation methods. NIST expresses CRM values as total concentrations, but USEPA methods use values based upon extractable or leachable concentrations. Because of this, certified NIST values for solid CRM may not be used directly to determine acceptance limits of PE samples. However, NIST values do not pose a problem with analysis of water PE samples, where extractable or leachable concentrations approximate NIST certified total concentrations.
- Base acceptance criteria on extractable or leachable concentrations.

4.5. Handling and Storage of PE Samples.

Guidelines for storage of fortified and real-world PE samples are given in this section as well as packing and shipping considerations.

4.5.1. Fortified PE Samples. All PE samples should be handled and stored with extreme care to ensure stability, integrity, purity, and authenticity. Follow these storage guidelines:

- Preserve and store PE samples with short shelf-lives (most fortified PE samples) like normal field samples. (Most of these samples are prepared on as needed basis.)
- Use the same sample containers required by the analytical methods for storage. Containers are selected for their inertness to the contents and their ability to prevent sample loss.
- Store PE samples for organic analyses in amber glass containers to avoid leaching of plasticizers from plastic containers or loss of target analytes due to absorption by plastic containers. Amber glass is recommended since some analytes are light sensitive.
- Use plastic bottles for metals to avoid leaching of trace metallic impurities from glass containers. Phenolic caps should not be used to avoid potential sample contaminations.
- Do not group analytes requiring different preservatives together in the same sample container.

- Fill the bottles for volatile organic samples completely to prevent loss of volatiles.

4.5.2. Real-World PE Samples. Although real-world PE samples used by USACE are generally very stable with long shelf-lives, the homogenized soils should be stored in a cool, dark, and dry place to retard degradation. Any inert and hermetically sealable containers are acceptable for storage. If needed, remove influences of laboratory humidity by conditioning the homogenized soils with calcium sulfate desiccant.

4.5.3. Packing and Shipping. Follow these guidelines for packing and shipping to ensure the integrity and quality of the PE samples:

- Use appropriate packaging materials, documentation, and shipping labels for all PE samples according to USACE, USEPA, and DOT regulations and guidelines.
- Use special precautions on labeling and packing double blind PE samples make them indistinguishable from regular field samples. Use packaging and containers identical to those used by field personnel sending the same types of field samples.
- Ship PE samples with short holding times immediately after preparation to allow adequate time for the contract laboratory to prepare and analyze them.
- Use overnight delivery services for shipping of all PE samples.

Chapter 5
Certification of PE Samples

PE samples are used to judge the performance of laboratories and data quality; therefore, PE samples must be certified on the basis of the following:

- Determination of the reference values. The reference values of PE samples are concentration values assigned to target analytes deemed to be within statistical limits. They must be based on scientifically valid and legally defensible procedures. Reference concentrations should be close to the true or prepared concentrations and are usually determined as the mean measured concentrations by a group of laboratories under control conditions.
- Determination of acceptance ranges of the reference concentrations based on symmetrical statistical prediction intervals around the mean measured concentrations.

PE samples sent to each laboratory must be equivalent to the PE samples upon which the acceptance limits were determined. Prior to determination of acceptance limits, the homogeneity, stability, and reproducibility of PE samples must be determined and proven within limits to ensure equivalent samples.

Steps to ensure certified, equivalent PE samples are discussed in this chapter. Sections include the following:

- Initial sample verification
- Steps to ensure homogeneity, stability, and reproducibility
- Tests for homogeneity, stability, and reproducibility
- Determination of reference values and acceptance limits
- Use of hybrid approach
- Statistical analysis
- Other considerations
- Documentation
- Standard operating procedures
- Confidentiality and ethical considerations

5.1. Initial Verification of PE Sample Composition.

Before certification of PE samples, the PE sample composition must be verified. Initial verifications ensure that there are no gross errors in the PE sample production process and serve as a baseline for evaluation of the certification process. Use these guidelines to complete verification.

- Use several reliable laboratories to verify the composition of the PE samples with multiple definitive methods of different measurement principles in addition to the analytical methods being used for proficiency testing. Although the composition of fortified PE samples is well known, conduct initial verification to ensure that the mean measured concentrations of prepared values are within acceptance ranges of analytical errors (including preparation and measurement errors).
- Be sure methods used for verification analysis have a standard deviation of lower than $0.3\bullet$, where $\bullet$ is the target standard deviation for the proficiency test of concern.
- Determine the composition of real-world PE samples with appropriate qualitative and quantitative analyses besides the analytical methods being used for laboratory evaluation. The composition of real-world PE samples is therefore based on the mean value of measured concentrations.

5.2. Ensuring Homogeneity, Stability, and Reproducibility.

PE samples sent to individual laboratories must be equivalent and have to remain equivalent prior to the expiration of sample holding times.

5.2.1. General Guidelines. General principles to guarantee homogeneity, stability, and reproducibility include:

- Test for homogeneity, stability, and reproducibility both within and between production batches.
- Use any generally accepted test procedures to ensure the consistency of analytes in each PE sample across the production run and through the life span of use at a 95% confidence level. Do not use PE samples that failed to pass the testing criteria in the USACE PE Program. See Section 5.3 for tests of homogeneity, stability, and reproducibility that can be performed by one or more reliable laboratories.
- Maintain proper documentation as evidence that PE samples are sufficiently homogeneous, stable, and reproducible.

The remainder of this section gives guidelines for ensuring equivalent samples for fortified aqueous, fortified soil, and real-world PE samples.

5.2.2. Fortified Aqueous PE Samples. Due to the nature of aqueous solutions, homogeneity is not a concern. However, the following steps are necessary to ensure stability and reproducibility of fortified aqueous PE samples. Additionally, stability and homogeneity of spiking reagents and solutions must be monitored.

5.2.2.1. General Guidelines.

- Spike the chemical standard solutions directly into the sample containers where loss of sample components to the walls of the glassware may be a problem.
- Designate experienced personnel to use dedicated syringes, pipetters, and glassware to minimize variability and maintain reproducibility between individually prepared PE samples.

5.2.2.2. Stability.

- Prepare and preserve PE samples properly to maintain stability within method-specified sample holding times.
- Check, at a minimum, stability of mean analyte concentration at the end of sample holding times. Inorganic analytes should be within ±10% of the initial mean concentrations; organic analytes should be ±15%.
- Select clean, inert containers, and consider chemical compatibility for selection of containers, starting materials, and stock solutions.
- Use ASTM Standard D4515 to determine the expiration date of PE samples under specific storage conditions recommended by the USEPA.

5.2.2.3. Spiking Standards.

- Check the homogeneity and stability of the spiking reagents and solutions on a routine basis. For liquid spiking solutions, warm the solutions to ambient temperature and vigorously shake to ensure a homogeneous mixture without precipitation.
- Prepare, verify, and use a new spiking reagent or solution if necessary.

5.2.3. Fortified Soil PE Samples. Homogeneity, stability, and reproducibility are all major concerns for fortified soil PE samples. Soil PE samples for explosives and volatile organic compounds are prepared by fortifying pre-weighed clean soils in vials and ampules, respectively. The entire amount of soil PE sample in a vial or ampule is used for each proficiency testing so that homogeneity in a vial or ampule is not a concern, but reproducibility of individual PE samples is a major concern. Use the following steps to ensure reproducibility of fortified soil PE samples. Additionally, stability and homogeneity of spiking reagents and solutions must be monitored.

5.2.3.1. General Guidelines. Follow these steps to ensure reproducibility:

- Grind solid spiking reagents into fine particles of similar size to that of sample matrices if individual soil PE samples are subsampled from a large quantity of fortified soil PE bulk materials.
- Use dedicated equipment and experienced personnel to ensure the reproducibility within and between batches.

5.2.3.2. Spiking Standards.

- Check the stability of the fortified soil PE samples and the spiking reagents on a routine basis and prior to use to ensure that no substantial changes in analyte concentrations have occurred. The frequency of stability checks depends on the characteristics of PE samples and reagents and the storage conditions.
- Refer to control charts and trend tests of historical data to determine significant degradation of analyte concentrations.

5.2.4. Real-World PE Samples. USACE routinely uses real-world soils and sediments to prepare real-world PE samples. Homogeneity and stability are major concerns for real-world PE samples. Sample reproducibility depends on homogeneity and stability of the bulk sample materials. Use these guidelines to ensure homogeneity:

5.2.4.1. PE Sample Homogeneity.

- Use PE sample materials of sufficient quantity and homogeneity to ensure representative subsamples with low sampling errors. (See Appendix D for more information on sampling error.)
- Check the homogeneity of PE materials by analysis of replicate PE samples from different sections of the bulk PE sample materials. The results should show no significant differences in concentrations for the replicates analyzed. A soil PE sample must be homogeneous to such a degree that residual difference between the compositions of PE samples will contribute virtually nothing to the variability of the results of participant laboratories.

5.2.4.2. PE Sample Stability. Most real-world soil or sediment PE samples are very stable under proper storage conditions. Because real-world PE samples are usually used for an extended time period, the stability of the bulk PE sample materials should be monitored using these guidelines:

- Carry out stability studies at elevated temperatures to accelerate the rate of sample degradations and reduce time to obtain sufficient data.
- Establish an expiration date for real-world PE samples.
- Analyze bulk PE sample materials periodically against set criteria for acceptable stability.
- Control chart results to document the validity of accepted values and their control limits and to detect any temporal trend in measured concentrations as an early warning. See Section E.6.5 of Appendix E for an example trend test.

5.3. Testing for Homogeneity, Stability, and Reproducibility.

Experimental and statistical considerations for each of these areas are described below.

5.3.1. Homogeneity. Follow these guidelines for tests of homogeneity:

5.3.1.1. Experiment Design.

- Test homogeneity for each analyte in every type of PE sample material or test analytes known to be most sensitive to problems.
- Use judgmental sampling to collect samples where heterogeneity is expected. Use random sampling only when there is no expected or suspected heterogeneity.
- Select multiple (e.g., five or more) PE samples from each of the beginning, middle, and end of a production run or from each of the top, middle, and bottom sections of a bulk real-world PE sample material.
- Test samples in replicate (e.g., duplicate or triplicate) to yield a minimum of 30 tests. The number of samples taken and replicate determinations depends on the level of uncertainty and the budget. If a highly precise method is used for homogeneity determination, perform duplicate tests on each of eight to ten randomly selected samples. A highly precise method means that the repeatability standard deviation is less than 30% of the total standard deviation for the proficiency testing of concern. For a less precise method, perform triplicate tests on each of five samples in random order to avoid systematic time variations.

5.3.1.2. Statistical Analysis. Guidelines for statistical analysis include:

- Use a traditional analysis of variance (ANOVA) F test to test the significance of the between-sample component of variances. If the F test is not significant at • = 0.05 level, the PE samples may be considered homogeneous. (See Section E.5 of Appendix E for an example calculation of a homogeneity test.)
- If the F test is significant, perform an additional test to assure that the significance represents a difference that could truly affect the evaluation of proficiency testing. The test compares the total standard deviation (include both between- and within-sample standard deviations) with the acceptance limits (i.e., 3•) of the PE sample. If the total standard deviation is less than 10% of the acceptance limit of target analyte, the PE sample can still be considered sufficiently homogeneous and used in the USACE PE Program.
- Where the results show the material to be not sufficiently homogeneous, reprocess the material (i.e., additional grinding and mixing), recheck the homogeneity, or select an alternative material. Another option is to relax the target standard deviation for that particular material to account for the variance the material will contribute to individual results. Except in minor excursions from sufficient

homogeneity, use such a practice with great care because it would destroy the utility of the proficiency test and the confidence of participant laboratories.

5.3.2. Stability. Stability is discussed for fortified and real-world PE samples along with guidelines for sample handling and spiking standards.

5.3.2.1. Fortified PE Samples. For fortified PE samples, follow these guidelines for stability testing:

- Use multiple samples (e.g., at least five) randomly selected from the production run.
- Conduct tests in duplicate at the end of the sample holding times as specified by the analytical methods used for laboratory evaluation.
- Compare results with the results of the initial verification tests. The means of the analytical results should not be statistically different at • = 0.05 level using a conventional *t* test. See Section E.6 of Appendix E for an example calculation of a stability test. If the means are significantly different, compare the difference with the size of the corresponding acceptance limits. If the difference between the means is less than 10% of the acceptance limits (i.e., 3•) of PE samples, the PE samples are considered sufficiently stable for the specific analytes.
- Retain and test one PE sample for each batch of PE samples prepared for samples with established shelf-lives.
- Control chart the results of these tests to document the continual validity of the sample stability.

5.3.2.2. Real-World PE Samples. For real-world PE samples, follow these guidelines:

- Conduct stability testing over an extended time period to cover the anticipated shelf-life.
- Conduct the initial test at the end of the method-specified sample holding time to ensure that there is no significant change in the composition of the PE sample. Thereafter, conduct tests at fixed intervals to assure long-term stability until the supply of the PE samples is exhausted.
- Use an experiment design similar to that of fortified PE samples.
- Use control charts to detect any trend in measured concentrations over time.

5.3.2.3. Sample Handling. The instability of PE samples may be caused by packing and shipping processes. Therefore, pack and mail PE samples for stability testing to the PE Sample Provider in the same way as to participant laboratories.

5.3.2.4. Spiking Standards. Test stability for spiking standards using these steps:

- Verify purity of prepared or purchased spiking reagents or solutions prior to the first use.
- Check the concentration of spiking reagents or solutions before each use and on a quarterly basis with appropriate methods.

- Use only spiking reagents or solutions with mean measured concentrations within ±10% (for inorganic) and ±15% (for organic) of the prepared values or the initial mean measured concentrations with a conventional two-sample *t* test at • = 0.05 level. Otherwise, discard the spiking reagents or solutions.
- Use control charts to monitor the stability and trend of degradation of spiking reagents or solutions over an extended time period.

5.3.3. Reproducibility. Although dedicated syringes, pipets, glassware, and personnel will be used to minimize the variations of PE samples produced within and between batches, check reproducibility of PE samples on a routine basis to ensure that equivalent PE samples are continually produced. Because certain PE samples of short holding times are prepared on an as needed basis, evaluate both within- and between-batch reproducibilities.

5.3.3.1. Within-Batch Testing. Follow these steps for within-batch reproducibility testing:

- Perform tests of multiple PE samples (e.g., at least five or more) randomly selected from a production batch.
- Test each sample in duplicate or triplicate, yielding a minimum of 15 tests.
- Use a traditional ANOVA F test to test the significance of the within-batch component of variances as for homogeneity tests. If the F test is not significant at the • = 0.05 level, the PE samples may be considered equivalent and the production process reproducible.
- If the F test is significant, conduct an additional test to assure that the significance represents a difference that truly could affect the evaluation of the results. To do this, compare the size of total within-batch variances with the acceptance limits of PE samples. If the total within-batch standard deviation is less than 10% of the acceptance limits (i.e., 3•) of the PE samples, the within-batch PE samples are considered sufficiently reproducible.

5.3.3.2. Between-Batch Testing.

- Select by random multiple (e.g., at least five or more) PE samples from different production batches.
- Use similar procedures to the within-batch reproducibility testing.

5.4. Determination of Reference Values and Acceptance Limits.

5.4.1. Determination of Reference Values. A reference value is usually derived from:

- A theoretical or established value based on scientific principles.
- An assigned or certified value based on experimental work of a certain national or international organization.

- A consensus or certified value based on collaborative experimental work under the auspices of a scientific or engineering group.
- The mean of a specified population of measurements if the above items are unavailable.

Reference values of PE samples are generally determined from the mean measured values, which should be close to the true or prepared values. When reference values are determined from the mean measured values, they reflect the recovery of target analytes or the bias of a specific analytical method. Therefore, reference values of PE samples might not be the same as true or prepared values.

5.4.2. Determination of Acceptance Limits. The acceptance limits of PE samples are determined by prediction intervals. These intervals are based on the statistical uncertainties from the means measured or reported values from round-robin testings. Test the reference values and the acceptance limits of PE samples meticulously both internally and externally prior to and during proficiency testing. See Section E.9 of Appendix E for an example determination of reference value and acceptance limits based on data from a round-robin testing.

5.4.3. Common Approaches. Several other approaches may also be used to determine the reference values and acceptance limits of PE samples. Based on technical and economical considerations, the common approaches of the USACE PE Program are listed below in the order of declining preference.

- Method performance data analysis
- Referee laboratory analysis
- Round-robin testing
- Error propagation analysis

These approaches are often used concurrently or sequentially, depending on PE sample requirements and available resources and technical information. The rest of this section explains each approach.

5.4.3.1. Method Performance Data Analysis. If reliable, published method performance data are available, the reference values and the acceptance limits of PE samples can be determined from method performance data published in literature or PE studies such as USEPA Water Supply (WS)/Water Pollution (WP) Programs. Method performance data are typically presented as linear regression equations that relate true or prepared concentrations to mean measured concentrations and standard deviations. Determine values and limits by following these steps:

- Set acceptance limits using the predicted means and standard deviations. Linear regression equations may only be used for prepared values that fall within the linear range of the method or the results might be biased.

- Base the warning and acceptance limits on 95% and 99% prediction intervals, respectively. (See Sections 5.4.2 and 6.1.6 for details on determination and use of warning and acceptance limits, respectively.)
- Make sure the method and matrix to be used in proficiency testing match those used to develop the method performance data to ensure reliability.
- Be sure concentrations of target analytes in the PE samples fall within the same concentration ranges as method performance data.

5.4.3.2. Referee Laboratory Analysis. This section describes qualifications of referee laboratories, how to use them, adjusting acceptance limits, analytical requirements, and establishing acceptance limits.

5.4.3.2.1. Qualification of Referee Laboratories. If reliable, published method performance data are not available, referee laboratory analysis shall be used to establish the reference values and acceptance limits. These laboratories use the mean measured concentration and the 95% and 99% prediction intervals to establish the reference values and the warning and acceptance limits of PE samples. Qualifications of referee laboratories include:

- Reliable with high performance standards.
- Validated and approved by the Program Manager according to the procedures described in USACE EM 200-1-1.
- Follow stipulated methods to prepare and analyze PE samples.

5.4.3.2.2. Using Reference Laboratories. Follow these steps to use referee laboratories:

- Send PE samples to referee laboratories for characterization before shipping to participating laboratories. However, send PE samples with short holding times to participating laboratories and a minimum of four referee laboratories at the same time.
- Investigate excessive high or low recoveries. Measured concentrations should be within 10% of the prepared concentration for the majority of target analytes. Analytes of poor recoveries and low concentrations may exceed 10%, but this may be acceptable on a case specific basis.
- Resubmit one or two PE samples periodically to check any potential degradation or temporal trends in analyte concentrations.

5.4.3.2.3. Adjusting Acceptance Limits. Because of the high performance of referee laboratories, their acceptance limits could be too tight. Adjust for this by:

- Ask referee laboratories to use multiple chemists, each using his/her own instruments, standards, reagents, etc., to prepare and analyze PE samples to simulate typical laboratory performance. If

use of a common instrument cannot be avoided, each chemist should establish new calibration curves to simulate independent analyses.
- Consider double blind PE samples to further reduce experimental bias, especially for the standard deviation.

5.4.3.2.4. Analysis Requirements. Guidelines for referee laboratories include the following:

- Ten replicate samples must be available from referee laboratory analysis for establishing acceptance limits.
- Sample extractions or digestions must be performed singly on separate days by one of several chemists or technicians.
- Each sample set must consist of one PE sample and accompanying QC samples, which include a method blank, a blank spike, a blank spike duplicate, and an independent reference sample.
- Samples must be analyzed by multiple operators on multiple instruments utilizing individually prepared calibration curves.
- QC acceptance criteria are method-specific and must be used to evaluate the recovery of analytes, not to reject data.
- All data meeting the minimal method-specific calibration criteria shall be used to calculate initial acceptance limits.
- Outliers that are not sustained by scientific reasoning or technical evidence shall not be removed from the data set before the standard deviations are calculated.

5.4.3.2.5. Establishing Acceptance Limits. To establish acceptance limits, follow these guidelines:

- For initial acceptance limits based on a few referee laboratory analyses, adopt the 99% and 99.9% confidence intervals around the mean measured concentration as the initial warning and control limits, respectively.
- When more data from participating laboratories are available, update the initial acceptance limits based on the mean measured concentration and the 95% and 99% prediction intervals of the pooled data.

5.4.3.3. Round-Robin Testing. Similar to referee laboratory analysis, in round-robin testing PE samples are analyzed by peer laboratories under control conditions. Data are combined to form a consensus. The number of participating laboratories is usually much larger than that of referee laboratory analysis, and the performances of the participating laboratories are unknown. The reference values and acceptance limits are determined based on the mean reported values and the associated prediction intervals. For example, the true values of target analytes in real-world PE samples are usually unknown so that the mean reported values from a round-robin testing are usually considered as the true values and are used as the reference values. Use normal 95% and 99% prediction intervals around the mean reported values as the warning and control limits, respectively. Due to complicated

sample matrices, the mean measured value and the acceptance limits for each target analyte in real-world PE samples are often matrix- and method-specific. Follow these guidelines for round-robin testing:

- Use it to verify the appropriateness of existing reference values and acceptance limits that are established with other approaches.
- Use other approaches of expectantly higher reliability, such as referee laboratory analysis, to verify the results of round-robin testing.
- Use round-robin testing or referee laboratory analysis to determine reference values and acceptance limits of fortified soil PE samples. It is difficult to determine these values using other methods because analytes' leachable levels must be known, and analytes in the spiking solution react with a solid matrix.

5.4.3.4. Error Propagation Analysis. Acceptance limits may also be calculated from error propagation analysis. Sometimes calculation is the only way to determine the true values and acceptance limits; it is very accurate and straightforward for fortified PE samples. The acceptance limits of fortified aqueous PE samples can be determined through error analyses of the sample preparation and analysis steps. Error propagation rules are used as guidelines to estimate determinate and indeterminate errors that would be experienced by the laboratory being evaluated. Indeterminate errors are always judgment calls and would be based on experience. The Factor-2 criterion (i.e., indeterminate errors are approximately equal to two times determinate errors) may be used as a good approximation for inclusion of indeterminate errors. The result is a relative error that can be multiplied by the expected target concentration of each analyte to obtain acceptance limits. Follow these guidelines for error propagation analysis:

- Use several referee laboratories and consensus values to characterize PE samples if biases are known to exist but cannot be accounted for.
- Ask PE Sample Providers to analyze four replicate samples from each batch of PE samples to confirm accuracy and precision if initial control limits are based on error propagation analysis or fixed percentages of true values.
- Ensure that the correct number of significant figures is retained during error propagation analysis.

5.5. Hybrid Approach.

USACE procedures for determining acceptance limits vary depending on the type of PE samples and the method performance information available. The preferred choice is published method performance data based on round-robin testings. If method performance data are unavailable, the USACE adopts a hybrid approach of referee laboratory analysis and round-robin testing. Steps involved in this approach include the following:

- Use the mean measured value of the results of referee laboratories to establish reference values.
- Use the 99% and 99.9% confidence intervals around the mean measured value to establish initial acceptance limits.
- Send the new PE sample along with a developed PE sample to participating laboratories and collect results for statistical analysis.
- Pool all data when a minimum twenty data points have been returned including those of referee laboratories.
- Reanalyze statistically to establish revised reference values and acceptance limits. Base the revised acceptance limits on 95% and 99% prediction intervals of the pooled data. They should be similar to the initial ones.
- Investigate possible sources of errors if the revised reference values and acceptance limits are substantially different from the initial ones.
- Statistically analyze PE results from peer laboratories on a regular basis to check for significant variations in bias and precision of PE sample analysis.
- Readjust the reference values and acceptance limits of PE samples one more time based on additional (e.g., a minimum of forty data points) new or pooled data.
- Monitor acceptance limits on a routine basis and adjust on an annual basis if needed.

5.6. Statistical Analysis.

General considerations for statistical analysis as well guidelines for distribution and outlier tests are given in this section.

5.6.1. General Considerations. Before applying any statistical techniques to analytical data, a preliminary data review is conducted to determine whether the data support underlying assumptions or if data modifications are necessary before further statistical analysis. The preliminary data review typically includes:

- Calculations of some basic statistical quantities including number of observations, measures of central tendency, dispersion, and distribution symmetry. These are useful for making inferences concerning the population used as a basis for the data.
- Graphical representations which are used to identify patterns and relationships within the data, confirm or disprove a hypothesis, and identify potential problems. This review reveals the structure of the data and identifies appropriate approaches and limitations of data use.

5.6.2. Distribution Tests. Chemical data can usually be analyzed using normal or Gaussian statistics. Acceptance limits for PE samples are determined from data sets that show normal distributions. Data can be tested for normal distribution using graphical representations such as normal probability plots, frequency plots, or histograms.

- Use a normal probability plot to determine whether data have a normal distribution, especially if a non-normal distribution is suspected.
- Use log-transformed data to test for normal distribution if the analytical method is highly variable and shows negative control limits for analyte recoveries.
- Test for the presence of outliers in a data set (see Section 5.6.3). Acceptance limits should not be determined from data sets that are strongly influenced by outliers.
- Use a larger population of samples or other statistical techniques such as outlier removal, non-parametric statistics, etc. to determine acceptance limits.

5.6.3. Outlier Tests. Outliers are measurements that are extremely large or small relative to the rest of the data and, therefore, are suspected of misrepresenting the population. Statistical outlier tests give evidence that an extreme value does not fit the distribution of the remainder of the data. Follow these steps for outlier testing:

- Use graphical representations such as frequency plots, histograms, normal probability plots, etc., to identify possible outliers.
- Evaluate suspected outliers with statistical tests such as Dixon's test, Grubbs' test, Cochran's test, Rosner's test, Walsh's test, or Youden Ranking test. (Refer to a statistical reference book or statistician to determine the proper test to use based on the data acquired.)
- Correct, retain, or discard outliers. The final judgment of whether an outlier is discarded or retained depends on the scientific judgment of the analyst and not just the outcome of the outlier test. Extreme values may be justifiably retained in a data set if other evidence outweighs the outlier test result.
- Discard an outlier value if there is a defensible explanation.
- Retain outlier if no good scientific judgment for it exclusion is found.
- Perform statistical analysis of the data set (e.g., mean and standard deviation) on both the full (outlier included) and the truncated (outlier excluded) data sets to measure the effect of the removed outlier.
- Discard scientifically indefensible outliers if the outlier exerts a large influence on the data set (e.g., it creates a large standard deviation or produces a large change in the calculated mean such that prediction intervals are useless).
- Document all statistical tests as well as the scientific judgments for discarding or retaining questionable data points.
- Use a minimum of 15 data points after rejection of outliers to calculate acceptance limits.

5.7. Other Considerations.

Other guidelines for ensuring the quality of PE samples given in this section include data accuracy, data comparability, program-wide statistical results, and PE Sample Provider requirements.

5.7.1. Data Accuracy. Accuracy of analyte concentrations in real-world PE samples depends on the amount of data available or affordable. To determine the composition of real-world PE samples, send split PE samples to a minimum of four referee laboratories. The consensus value from several referee laboratories should be a better estimate of true value than any single measurement.

5.7.2. Data Comparability. All PE samples must be analyzed with the same methodology by participating laboratories as well as the referee laboratories. Any deviations from the standard methods may make data not comparable.

- Use the results of all PE sample analyses to develop control charts displaying the true concentration, recovery ranges, and bias for each target analyte. Control charts will help detect trends and prevent problems.

5.7.3. Program-Wide Statistical Results. Compile program-wide statistical results using these guidelines:

- Analyze the PE sample results produced by all contract laboratories on a regular basis.
- Document the mean values and the associated uncertainties of target analytes.
- Use results to adjust the acceptance limits in order to observe the relative performance of each laboratory using a given protocol against its peers. The USACE may adjust the acceptance limits on any given PE sample to compensate for unanticipated difficulties with a particular sample or analysis.

5.7.4. PE Sample Provider Requirements. Each PE Sample Provider is required to have in place within its own dedicated facility the capability to design, produce, test, and distribute PE samples, and to provide data analysis and reporting functions for any series of PE samples. The technical staff, instrumentation, and computer capabilities must be able to support these tasks.

5.8. Documentation.

PE Sample Providers shall establish and maintain procedures to control all documents and data that relate to the design, development, production, certification, and use of PE samples. The documentation should provide clear evidence that the PE samples are developed and used with scientific and legal defensibility. The documentation should cover, but not be limited to, the following areas.

- Protocols for production of PE samples.
- Certification of PE samples with validated definitive methods by referee laboratories and/or collaborative trials.
- Demonstration of statistical control in production processes with fully established control charts.
- Demonstration of data traceability with scientific and legal defensibility.

Follow ISO Guides 31, 34, and 35 to document the preparation and contents of PE samples. At a minimum, document the following:

- PE sample name
- Unique identification number and batch number
- Description of the PE sample
- Certified compositions and uncertainties
- Intended use of the PE sample
- Instructions for the correct use of the PE sample
- Sources of the PE materials
- Preparers of the PE samples
- Name and address of the certification organization
- Stability, transportation, and storage instructions
- Preparation method of the PE sample
- Certification method of the PE sample
- Names of participating laboratories
- Name of certification officer
- Date of certification

5.9. Standard Operating Procedures.

In order to prepare reliable and reproducible PE samples, it is imperative that the PE Sample Providers use SOPs. Use USEPA document QA/G-6, "Guidance for the Preparation of Standard Operating Procedures (SOPs) for Quality-Related Documents," to prepare SOPs. Follow these guidelines for writing PE sample SOPs:

- Include actual procedures used in the laboratory to ensure that reproducible results can be achieved by following SOPs.
- Prepare SOPs as part of the planning process and complete before PE sample preparation work begins.
- Submit to the Program Manager for review and approval before production.
- Avoid ambiguous statements like "air dried at ambient temperature" or "1:10 dilution" when "18 to 22°C" or "tenfold dilution" is meant. (The "1:10" may be confused with one part concentrate diluted with ten parts of diluent, which is really an 11-fold dilution.)
- Note changes and observations during preparation for subsequent SOP revisions.

5.10. Confidentiality and Ethical Considerations.

Concerns of confidentiality and ethics for PE Sample Providers and participating laboratories are given below.

5.10.1. PE Sample Providers. Concerns include the following:

- Data generated from PE sample analysis is confidential.
- No portion of the production, testing, distribution, data collection, or data reporting may be subcontracted by PE Sample Providers.
- PE Sample Providers shall declare that they do not have potential conflict of interest with any laboratory seeking, or having, the USACE environmental laboratory validation. PE Sample Providers shall notify the Program Manager of any actual or potential organizational conflicts of interest including, but not limited to, financial interest, sharing of personnel, facility, or instrumentation with any laboratories.
- The reference values and acceptance limits of PE samples are proprietary information of the USACE and shall not be disclosed by PE Sample Providers without written approval of the Program Manager.
- PE Sample Providers shall not sell, distribute, or provide PE samples of similar or identical design and concentration to any laboratories who may, will, or are seeking USACE environmental laboratory validation.

5.10.2. Participating Laboratories. Guidelines regarding confidentiality include the following:

- The confidentiality of laboratories participating in the USACE PE Program will be maintained. The identity of participants should be treated as confidential information with limited access, and this shall extend to any subsequent remedial advice or actions applied to a laboratory exhibiting poor performance.
- Participating laboratories shall not send any PE samples or a portion of any PE samples to another laboratory for any analysis, shall not exchange information with another laboratory concerning any PE samples, and shall not attempt to obtain the target values of any PE samples from PE Sample Providers.

Chapter 6
Use of PE Samples

Project-specific PE samples are key tools for evaluating a laboratory's performance of sample analysis. PE samples are used to assess turnaround, customer service, and report content as well as data accuracy. Refer to EM 200-1-1, "Validation of Analytical Chemistry laboratories," for specifics regarding the USACE Laboratory Validation Program.

The uses of PE samples discussed in this chapter include:

- Pre-contract proficiency testing
- Post-contract performance monitoring
- Compliance and corrective actions

6.1. Pre-Contract Proficiency Testing.

6.1.1. Overview. The USACE laboratory validation process for HTRW environmental projects consists of five major sequential steps:

(1) Review of qualification documents
(2) Proficiency testing and/or evaluation of SOPs
(3) On-site laboratory inspection
(4) Resolution of inspection findings
(5) Decision of validation status

The laboratory validation procedures described in this section relate directly to PE sample use and include:

- Initiation of laboratory validation
- Initial PE samples
- Analysis of PE samples
- Reporting of PE sample results
- Evaluation of PE sample results
- Tentatively Identified Compounds (TIC)
- Requested corrections
- Unavailable PE samples
- Remedial PE samples
- Multimedia validation

6.1.2. Initiation of Laboratory Validation. A written request for evaluation of laboratory performance from a customer to the Program Manager begins the process. Upon receipt of a

request, the Program Manager informs the laboratory of the forthcoming laboratory validation, sends an information package about the laboratory validation process and a questionnaire to the laboratory, and requests copies of laboratory's qualification documents. The qualification documents, such as QA Manual, SOPs, etc., shall provide pertinent information for the Program Manager to determine whether the laboratory is capable of meeting project requirements and whether project-specific PE samples should be sent.

6.1.3. Initial PE Samples. PE samples shall be prepared and sent out by designated PE Sample Providers through overnight delivery service. All PE samples shall be preserved and shipped according to USACE, USEPA, and DOT regulations or guidelines. A full chain-of-custody (COC) shall be maintained for each shipment of PE samples. Sample-specific instructions for individual PE samples are included in the shipment. (General guidelines for analysis and reporting are sent in the initial package from the Program Manager.)

6.1.4. Analysis of PE Samples. The following sections describe selection of PE samples, metal PE samples, analytical methods, modification and approval requirements of analytical methods, and laboratory analysis.

6.1.4.1. Selection of PE Samples. PE samples designed for project-required methods will be prepared and shipped to participating laboratories. If no PE sample for project-required method is available from the USACE in a timely manner, the Program Manager may consider a replacement PE sample. The replacement of PE samples is allowed only if both analyses are similar in technical nature, and the laboratory performs both analyses in-house on a routine basis.

6.1.4.2. Metal PE Samples. For metal analysis, the laboratory validation is granted based on a combination of the analytical methods used and the number of metal elements in the PE samples passed. The analytical methods are usually classified as FLAA, GHAA, CVAA, GFAA, ICP-AES, or ICP-MS. The metal elements are grouped into the following four categories.

- Category I: Eight Resource Conservation and Recovery Act (RCRA) metal elements including arsenic, barium, cadmium, chromium, lead, mercury, selenium, and silver.
- Category II: Barium plus thirteen Priority Pollutant (PP) metal elements including antimony, arsenic, beryllium, cadmium, chromium, copper, lead, mercury, nickel, selenium, silver, thallium, and zinc.
- Category III: Twenty-three USEPA Contract Laboratory Program (CLP) Target Analyte List (TAL) metal elements including aluminum, antimony, arsenic, barium, beryllium, cadmium, calcium, chromium, cobalt, copper, iron, lead, magnesium, manganese, mercury, nickel, potassium, selenium, silver, sodium, thallium, vanadium, and zinc.
- Category IV: Any other combinations of metal elements.

Based on project requirements for metal analysis, one of the above categories of metal PE samples in combination with the analytical method(s) will be selected for proficiency testing. A laboratory may volunteer for any one of the four categories of metal PE samples provided the minimum project-required metals are analyzed.

6.1.4.3. Analytical Methods. Standard analytical methods from the following sources are usually required for USACE's HTRW environmental projects and hence for the analysis of PE samples.

- *Test Methods for Evaluating Solid Waste*, USEPA SW-846.
- *Methods for Chemical Analysis of Water and Wastes*, USEPA-600/4-79-020.
- *Methods for the Determination of Organic Compounds in Drinking Water*, USEPA-600/4-88/039.
- *Statements of Work for Organics Analysis, Inorganics Analysis, and Dioxin Analysis*, USEPA Contract Laboratory Program.
- Other published, standard methods of the most recent versions from USEPA, American Society for Testing and Materials, American Public Health Association, American Water Works Association, Water Environment Federation, United States Geological Survey, National Institute for Occupational Safety and Health, Department of Energy, and other equally qualified agencies or organizations.

6.1.4.4. Modification of Analytical Methods. A laboratory must analyze PE samples using the preparation and analytical methods specified on the chain-of-custody form enclosed in the sample shipping container. Any changes or modifications in preparation or analytical methods of PE samples must be approved by the Program Manager. Use of nonstandard or modified standard methods without preapproval from the Program Manager may result in failure of PE sample analysis.

6.1.4.5. Approval for Modified Analytical Methods. Because the acceptance limits of PE samples are developed for specific combinations of sample preparation and analytical methods, the acceptance limits may not be applicable if methods are modified. If a laboratory plans to use modified standard methods, it must submit its in-house method SOPs and method validation data (including MDL, initial performance demonstration, QC limits on precision and bias, chromatograms, etc.) to the Program Manager for approval prior to proficiency testing. For technical and/or cost reasons, PE samples for nonstandard or modified standard methods may not be available in a timely manner.

6.1.4.6. Laboratory Analysis. A laboratory shall treat PE samples as regular field samples and follow the specified methods and any sample-specific instructions to analyze and report PE samples. A laboratory shall also perform all method-required QC analyses that normally include, but are not limited to, analysis of blanks, spikes, duplicates, and QA samples. If the amount of a

PE sample provided is not enough for all QC analyses, the QC analyses shall be performed on spiked reagent water or clean solid matrices.

6.1.5. Reporting of PE Sample Results. Guidelines for reporting PE samples including QC and raw data, format and contents, and time requirements are described in this section.

6.1.5.1. General Reporting Requirements. A laboratory shall report the concentrations of all target analytes, specified by the PE Sample Providers, including estimated values, MDL, Method Quantitation Limits (MQL), Method Reporting Limits (MRL), and dilution factors. (Refer to EM 200-1-3 for the definition and use of MDL, MQL and MRL.) All data for soil/sediment PE samples shall be reported on a dry-weight basis along with percent moisture on an as-received basis, or per sample-specific instruction. No data shall be corrected for spike recoveries or blank contaminations.

6.1.5.2. QC and Raw Data. All batch-associated QC data including blank analysis, replicate analysis, spike recovery, etc., that are required by the analytical methods shall also be reported. Raw data, including sample preparation and run logs, instrument calibrations, chromatograms, calculations, etc., are generally not required but should be available if requested for review by the Program Manager. The data package of PE sample analysis shall be received by the Program Manager within 20 calendar days after receipt of the PE samples, unless otherwise instructed. For projects requiring quick turnarounds, the turnaround times of the PE samples shall also be reduced accordingly. Failure to analyze the PE samples correctly or within the required time frame may result in termination of the validation process.

6.1.5.3. Format and Contents. A laboratory may use its standard data package to report PE sample results, but the data package must be paginated and contain, at a minimum, the following information:

- A cover sheet containing the report's title and date and the laboratory's name and location.
- Table of contents.
- A case narrative including problems with PE sample analysis.
- Sample preparation information including the date and method of digestion, extraction, and cleanup procedures.
- Analytical results including the date and method of analysis, analyte concentrations, MDL, MQL, MRL, and dilution factors.
- Summary of the batch-associated QC data and acceptance limits to demonstrate data quality (precision and bias).
- Phone conversation records on major issues related to PE sample analysis.
- A chain-of-custody report.

6.1.5.4. Time Requirements. Failure to submit required information within the required time frame may result in termination of the validation process. It is the responsibility of the laboratory to keep the Program Manager informed early of any problems with PE sample analyses that would affect the return of results within a required time frame.

6.1.6. Evaluation of PE Sample Results. Acceptance limits and evaluation criteria of PE samples are discussed below.

6.1.6.1. Evaluation Report. Within ten calendar days after receipt of PE sample results, the PE Sample Provider shall prepare and send a written evaluation report to the Program Manager for review. The report shall contain the following information:

- Laboratory name, location (city and state), point of contact, phone and facsimile numbers, and e-mail address if available.
- Dates that PE samples were shipped and results were received.
- PE sample name, control number, and preparation and certification methods.
- Laboratory results, reference values, and warning and control limits of each target analyte.
- Narratives about special problems or issues.
- Follow-ups on failed parameters.
- A line-item summary report to present the status of proficiency testing. The summary report shall include the name of each PE sample, names of target analytes correctly identified but quantitated outside control limits, and the number of false positives and/or negatives reported for each PE sample without disclosure of their identities. More details on the format and contents of evaluation reports can be found in EM 200-1-1.

6.1.6.2. Status of Proficiency Testing. The Program Manager will review the evaluation report and determine the pass/fail status of each PE sample based on the statistically established acceptance limits of each PE sample and the method-specified acceptance criteria of the internal QC samples. The acceptance limits for each parameter/analyte are based on the 95% and 99% prediction intervals which are set as the warning and control limits, respectively.

6.1.6.3. Evaluation Criteria. Described in this section are the general evaluation rules and acceptance criteria for all, single-analyte, and multiple-analyte analyses; metal analysis by AA; and metal analysis by ICP. The number of target analytes is based on the target analytes in the PE samples instead of the number of analytes cited by the analytical methods.

6.1.6.3.1. All Chemical Analyses.

- All method-specific QC data are within project- or method-specified acceptance criteria.
- False positives or negatives are treated as outside control limits.

6.1.6.3.2. Single-Analyte Analyses.

- Each individual analyte is within control limits.

6.1.6.3.3. Multiple-Analyte Analyses. The number of target analytes that are allowed outside warning and control limits is based on the probability of pass/fail status of target analytes under a binomial distribution.

- For analyses with two to five target analytes: All target analytes are within warning limits with the exception of no more than two target analytes between warning and control limits.
- For analyses with six to fifteen target analytes: All target analytes are within warning limits with the exceptions of no more than two target analytes outside warning limits and no more than one of the two target analytes outside control limits.
- For analyses with sixteen to forty-five target analytes: All target analytes are within warning limits with the exceptions of no more than four target analytes outside warning limits and no more than two of the four target analytes outside control limits.
- For analyses with forty-six to eighty-five target analytes: All target analytes are within warning limits with the exceptions of no more than six target analytes outside warning limits and no more than three of the six target analytes outside control limits.

6.1.6.3.4. Metal Analysis by AA.

- All target analytes are within control limits.

6.1.6.3.5. Metal Analysis by ICP. The number of target analytes that are allowed outside warning and control limits are based on the probability of pass/fail status of target analytes under a binomial distribution.

- For analyses with two to five target analytes (e.g., Categories I and IV metals): All target analytes are within warning limits with the exception of no more than two target analytes between warning and control limits.
- For analyses with six to fifteen target analytes (e.g., Category II metals): All target analytes are within warning limits with the exceptions of no more than two target analytes outside warning limits and no more than one of the two target analytes outside control limits.
- For analyses with sixteen to thirty target analytes (e.g., Category III metals): All target analytes are within warning limits with the exceptions of no more than three target analytes outside warning limits and no more than one of the three target analytes outside control limits.

6.1.7. Tentatively Identified Compounds (TIC). Some non-target analytes may exist in the water or soil/sediment PE samples for volatile and semivolatile organic analyses. A laboratory

shall use NIST/USEPA/Mass Spectral Data Center (MSDC), or any other USEPA-approved mass spectral libraries to tentatively identify and quantify up to ten non-target volatile organic compounds and twenty non-target semivolatile organic compounds that exhibit the strongest ion current signals. These compounds must not be system monitoring compounds. Identification of these compounds, based on spectral interpretation procedures, will be evaluated and integrated into the evaluation process for volatile and semivolatile organic PE sample results.

6.1.8. Requested Corrections. The Program Manager will send a copy of the line-item summary report to the laboratory for information and/or any needed corrective actions. Due to confidentiality, the reference values and the warning and control limits of any batch of PE samples shall not be disclosed to laboratories until the batch is discontinued. A laboratory will be allowed to provide revised data for failed parameters if problems such as calculation or transcription errors can be identified. If a laboratory is requested by the Program Manager to check its proficiency testing results, the laboratory shall return revised results for failed parameters within five calendar days.

6.1.9. Unavailable PE Samples. For parameters without available PE samples (e.g., radioactivity, air toxics, petroleum hydrocarbons, etc.) from the USACE, the laboratory validation for these parameters will be based solely on the qualification documents of the laboratory. The documents shall include:

- Copies of laboratory QA manual, method SOPs, and in-house method performance data for MDL, precision, and bias. See EM 200-1-3 for SOP requirements. If an SOP is deemed unacceptable, the laboratory will have 20 calendar days to submit a revised SOP to the Program Manager for a second review.
- Laboratory certificates or licenses.
- The most recent two rounds of PE sample results from other government and/or private agencies.

6.1.10. Remedial PE Samples. After data and SOP revisions, a laboratory must pass, at a minimum, more than 50 percent of all project-required parameters within 50 calendar days from receipt of the first set of PE samples (or from request of SOPs if PE samples are not available). If it does not, the validation process will be terminated. The Program Manager will notify all affected customers for remedial actions immediately. After a laboratory passes more than 50 percent of all parameters, the Program Manager will contact the laboratory to schedule an on-site inspection within ten calendar days. Depending on the results of an on-site inspection, the Program Manager may send an additional set of PE samples for failed parameters or any parameters with major deficiencies noted during the on-site laboratory inspection.

6.1.11. Multimedia Validation. The majority of PE samples available from the USACE are in water and/or soil/sediment matrices. See the following table for validation type awarded based on PE sample type passed.

Table 6-1. Validation Parameters

Sample Type	Validation Type
Water PE samples only.	Multimedia.
Both water and soil/sediment PE samples.	Multimedia if both matrices are passed.
Both water and soil/sediment PE samples.	Water samples only if water PE samples pass and soil/sediment PE samples fail.
Both water and soil/sediment PE samples.	No validation in any matrix type if water PE samples fail.

6.2. Post-Contract Performance Monitoring.

The following section describes procedures for post-contract performance monitoring, possible follow-up actions, and special requests for proficiency testing.

6.2.1. Continual Performance Monitoring. In order to measure laboratory performance after the validation process, the Program Manager may send additional PE samples on a quarterly or as needed basis. This depends on the laboratory's past performance as well as the importance of the laboratory's USACE project. These quarterly PE samples may be either single or double blind and are shipped from the field or through a fictitious contract. The results will be evaluated based on the accuracy of analyte identification and quantitation, batch-associated QC data, turnaround time, content of reports, etc. Results from the analysis will be used by the Program Manager to monitor the laboratory's ongoing ability to produce acceptable analytical data.

6.2.2. Follow-Up Actions. The results of quarterly PE samples will be used to determine one of these follow-up actions:

- Acceptable and No Response Required: Data meets all of the evaluation criteria as previously described. No response is required.
- Acceptable but Response Explaining Deficiencies Required: Deficiencies exist in the laboratory performance. Within ten calendar days of receipt of notification from the Program Manager, the laboratory shall submit written response to describe the problems, corrective actions taken or to be taken, and supporting documentation including implementation

schedules. Based on the deficiencies and responses, the Program Manager may send additional PE samples or perform additional inspections to verify and evaluate corrective actions. If no response is received, the validation status of the laboratory may be suspended or revoked.

6.2.3. Special Requests for Proficiency Testing. Upon request from a customer(s), the Program Manager will send PE samples to evaluate laboratory performance during field sample analysis. The PE samples may be provided as single or double blind and can be specially designed to detect any suspected problems. If double blind PE samples are shipped, take following precautions to ensure the PE samples indistinguishable from regular field samples:

- Using the same bottles, labels, chain-of-custody forms, sample coolers, shipping location, etc., as those of actual environmental samples.
- Setting up a fictitious contract with the laboratory to be evaluated to hide the identity of double blind PE samples.

Customers and the Program Manager must coordinate closely to ensure the success of double blind PE samples.

6.3. Complaints and Corrective Actions.

The responsibilities of the PE Sample Providers and Program Manager for handling complaints are outlined below:

6.3.1. PE Sample Providers. PE Sample Providers shall establish, document, and maintain procedures for handling complaints from participating laboratories. Their specific responsibilities are to:

- Analyze the complaints to detect and eliminate potential causes of nonconforming PE samples.
- Notify the Program Manager of any complaints.
- Obtain approval for corrective actions
- Implement and document changes in procedures resulting from corrective actions.

6.3.2. Program Manager. The Program Manager shall investigate and resolve all complaints regarding the proficiency testing program within 20 calendar days from receipt of complaints. Other responsibilities include:

- Discontinuing use of questionable or problematic PE samples immediately until all problems are resolved and PE samples are recertified.
- Providing a yearly summary of all complaints received and resolutions taken to the HQUSACE.

Appendix A
References

A.1. Required Publications.

1. USACE ER 385-1-92. "Safety and Occupational Health Document Requirements for Hazardous, Toxic and Radioactive Waste (HTRW) and Ordnance and Explosive Waste (OEW) Activities." Washington, DC: United States Army Corps of Engineers, 1994.

2. USACE ER 1110-1-12. "Quality Management." Washington, DC: United States Army Corps of Engineers, 1993.

3. USACE ER 1110-1-263. "Chemical Data Quality Management for Hazardous, Toxic, Radioactive Waste Remedial Activities." Washington, DC: United States Army Corps of Engineers, 1998.

4. USACE EM 200-1-1. "Validation of Analytical Chemistry Laboratories." Washington, DC: United States Army Corps of Engineers, 1994.

5. USACE EM 200-1-3. "Requirements for the Preparation of Sampling and Analysis Plans." Washington, DC: United States Army Corps of Engineers, 1994.

6. USACE EM 200-1-6. "Chemical Quality Assurance for HTRW Projects." Washington, DC: United States Army Corps of Engineers, 1997.

7. USACE MFR. "Interim Chemical Data Quality Management (CDQM) Policy for USACE Hazardous, Toxic, and Radioactive Waste (HTRW) Projects." Washington, DC: United States Army Corps of Engineers, November 23, 1998.

8. ASTM D4841. "Standard Practice for Estimation of Holding Time for Water Samples Containing Organic and Inorganic Constituents." West Conshohocken, PA: American Society for Testing and Materials, 1988(1998).

9. ASTM D5522. "Standard Specification for Minimum Requirements for Laboratories Engaged in Chemical Analysis of Soil, Rock, and Contained Fluid." West Conshohocken, PA: American Society for Testing and Materials, 1999a.

10. ASTM D5810. "Standard Guide for Spiking into Aqueous Samples." West Conshohocken, PA: American Society for Testing and Materials, 1996.

11. ASTM D6323. "Standard Guide for Laboratory Subsampling of Media Related to Waste Management Activities." West Conshohocken, PA: American Society for Testing and Materials, 1998.

12. ASTM E178. "Standard Practice for Dealing with Outlying Observations." West Conshohocken, PA: American Society for Testing and Materials, 1994.

13. ASTM E826. "Standard Practice for Testing Homogeneity of Materials for Development of Reference Materials." West Conshohocken, PA: American Society for Testing and Materials, 1985(1996)• 1.

14. ASTM E1301. "Standard Guide for Proficiency Testing by Interlaboratory Comparisons." West Conshohocken, PA: American Society for Testing and Materials, 1995.

15. ASTM F1469. "Standard Guide for Conducting a Repeatability and Reproducibility Study on Test Equipment for Nondestructive Testing." West Conshohocken, PA: American Society for Testing and Materials, 1999.

16. Gy, Pierre M. "Sampling of Heterogeneous and Dynamic Material Systems: Theories of Heterogeneity, Sampling and Homogenizing (Data Handling in Science and Technology)." Oxford, England: Elsevier Science Publishers Ltd., 1992.

17. Hahn, Gerald, J. "Finding an Interval for the Next Observation from a Normal Distribution." *Journal of Quality Technology*, 1, 1969, pp. 168-171.

18. Hahn, Gerald, J. "Factors for Calculating Two-Sided Prediction Intervals for Samples from a Normal Distribution." *Journal of the American Statistical Association*, 64, 1969, pp. 878-888.

19. Hahn, Gerald, J. "Additional Factors for Calculating Prediction Intervals for Samples from a Normal Distribution." *Journal of the American Statistical Association*, 65, 1970, pp. 1668-1676.

20. Hahn, Gerald, J. "Statistical Intervals for a Normal Population, Part I. Tables, Examples and Applications." *Journal of Quality Technology*, 2, 1970, pp. 115-125.

21. Hahn, Gerald, J. "Statistical Intervals for a Normal Population, Part II. Formulas, Assumptions, Some Derivations." *Journal of Quality Technology*, 2, 1970, pp. 195-206.

22. ISO 2602. "Statistical Interpretation of Data - Estimation of the Mean - Confidence Interval." Geneva, Switzerland: International Organization for Standardization, 1980.

23. ISO 2854. "Statistical Interpretation of Data - Techniques of Estimation and Tests Relating to Means and Variances." Geneva, Switzerland: International Organization for Standardization, 1976.

24. ISO 3301. "Statistical Interpretation of Data - Comparison of Two Means in the Case of Paired Observations." Geneva, Switzerland: International Organization for Standardization, 1975.

25. ISO 3494. "Statistical Interpretation of Data - Power of Tests Relating to Means and Variances." Geneva, Switzerland: International Organization for Standardization, 1976.

26. ISO 5479. "Statistical Interpretation of Data - Tests for Departure from the Normal Distribution." Geneva, Switzerland: International Organization for Standardization, 1997.

27. ISO 5725-1. "Accuracy (Trueness and Precision) of Measurement Methods and Results - Part 1: General Principles and Definitions." Geneva, Switzerland: International Organization for Standardization, 1994.

28. ISO 5725-2. "Accuracy (Trueness and Precision) of Measurement Methods and Results - Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method." Geneva, Switzerland: International Organization for Standardization, 1994.

29. ISO 5725-3. "Accuracy (Trueness and Precision) of Measurement Methods and Results - Part 3: Intermediate Measures of the Precision of a Standard Measurement Method." Geneva, Switzerland: International Organization for Standardization, 1994.

30. ISO 5725-4. "Accuracy (Trueness and Precision) of Measurement Methods and Results - Part 4: Basic Methods for the Determination of the Trueness of a Standard Measurement Method." Geneva, Switzerland: International Organization for Standardization, 1994.

31. ISO 5725-5. "Accuracy (Trueness and Precision) of Measurement Methods and Results - Part 5: Alternative Methods for the Determination of the Precision of a Standard Measurement Method." Geneva, Switzerland: International Organization for Standardization, 1994.

32. ISO 5725-6. "Accuracy (Trueness and Precision) of Measurement Methods and Results - Part 6: Use in Practice of Accuracy Values." Geneva, Switzerland: International Organization for Standardization, 1994.

33. ISO 7870. "Control Charts - General Guide and Introduction." Geneva, Switzerland: International Organization for Standardization, 1993.

34. ISO 7966. "Acceptance Control Charts." Geneva, Switzerland: International Organization for Standardization, 1993.

35. ISO 8258. "Shewhart Control Charts." Geneva, Switzerland: International Organization for Standardization, 1991.

36. ISO 9001. "Quality Systems - Model for Quality Assurance in Design, Development, Production, Installation and Servicing." Geneva, Switzerland: International Organization for Standardization, 1994.

37. ISO 9002. "Quality Systems - Model for Quality Assurance in Production, Installation and Servicing." Geneva, Switzerland: International Organization for Standardization, 1994.

38. ISO 9003. "Quality Systems - Model for Quality Assurance in Final Inspection and Test." Geneva, Switzerland: International Organization for Standardization, 1994.

39. ISO 10011-1. "Guidelines for Auditing Quality Systems - Part 1: Auditing." Geneva, Switzerland: International Organization for Standardization, 1990.

40. ISO 10011-2. "Guidelines for Auditing Quality Systems - Part 2: Qualification Criteria for Quality Systems Auditors." Geneva, Switzerland: International Organization for Standardization, 1991.

41. ISO 10011-3. "Guidelines for Auditing Quality Systems - Part 3: Management of Audit Programs." Geneva, Switzerland: International Organization for Standardization, 1991.

42. ISO/IEC Guide 25. "General Requirements for the Competence of Calibration and Testing Laboratories." Geneva, Switzerland: International Organization for Standardization, 1990.

43. ISO/IEC Guide 28. "General Rules for a Model Third-Party Certification System for Products." Geneva, Switzerland: International Organization for Standardization, 1982.

44. ISO Guide 30. "Terms and Definitions Used in Connection with Reference Materials." Geneva, Switzerland: International Organization for Standardization, 1992.

45. ISO Guide 31. "Contents of Certificates of Reference Materials." Geneva, Switzerland: International Organization for Standardization, 1981.

46. ISO Guide 34. "Quality System Guidelines for the Production of Reference Materials." Geneva, Switzerland: International Organization for Standardization, 1996.

47. ISO Guide 35. "Certification of Reference Materials - General and Statistical Principles." Geneva, Switzerland: International Organization for Standardization, 1989.

48. ISO/IEC Guide 43-1. "Proficiency Testing by Interlaboratory Comparison - Part 1: Development and Operation of Proficiency Testing Schemes." Geneva, Switzerland: International Organization for Standardization, 1996.

49. ISO/IEC Guide 43-2. "Proficiency Testing by Interlaboratory Comparison - Part 2: Selection and Use of Proficiency Testing Schemes by Laboratory Accreditation Bodies." Geneva, Switzerland: International Organization for Standardization, 1996.

50. ISO/IEC Guide 56. "An Approach to the Review by a Certification Body of Its Own Internal Quality System." Geneva, Switzerland: International Organization for Standardization, 1989.

51. ISO/IEC Guide 65. "General Requirements for Bodies Operating Product Certification Systems." Geneva, Switzerland: International Organization for Standardization, 1996.

52. ISO/IEC Guide 17025. "General Requirements for the Competence of Testing and Calibration Laboratories." Geneva, Switzerland: International Organization for Standardization, 1999.

53. Lawn, Richard E., Michael Thompson, and Ronald F. Walker. "Proficiency Testing in Analytical Chemistry." Cambridge, England: The Royal Society of Chemistry, 1997.

54. Pitard, Francis F. "Pierre Gy's Sampling Theory and Sampling Practice - Heterogeneity, Sampling Correctness, and Statistical Process Control." Boca Raton, FL: CRC Press, 1993.

55. Thompson, Michael and Roger Wood. "The International Harmonized Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories." *Journal of AOAC International,* 76, 1993, pp. 926-940.

56. USACE CRREL Special Report 96-5. "Obtaining and Transferring Soils for In-Vial Analysis of Volatile Organic Compounds." Washington, DC: United States Army Corps of Engineers, 1996.

57. USEPA/600/R-97/139. "National Environmental Laboratory Accreditation Conference, Constitution, Bylaws and Standards." Washington, DC: United States Environmental Protection Agency, July 1997 and the most recent revision.

58. USEPA QA/G-9D. "Data Quality Evaluation Statistical Toolbox (DataQUEST)." Washington, DC: United States Environmental Protection Agency, December 1997.

59. Youden, W. J. and E. H. Steiner. "Statistical Manual of the Association of Official Analytical Chemists." Arlington, VA: Association of Official Analytical Chemists, 1975.

A.2. Related Publications.

1. 29 CFR 1910.1450. "Occupational Exposure to Hazardous Chemicals in Laboratories."

2. ANSI/ASQC E4. "Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs." Milwaukee, WI: American Society for Quality Control, 1994.

3. ANSI/NCSL Z540-1. "Calibration Laboratories and Measuring and Test Equipment - General Requirements." Boulder, CO: National Conference of Standards Laboratories, 1994.

4. ANSI/NCSL Z540-2. "Guide to the Expression of Uncertainty in Measurement." Boulder, CO: National Conference of Standards Laboratories, 1997.

5. ASTM Special Technical Publication 355. "Manual for Conducting an Interlaboratory Study of a Test Method." West Conshohocken, PA: American Society for Testing and Materials, 1963.

6. ASTM Special Technical Publication 845. "Statistics in the Environmental Science." West Conshohocken, PA: American Society for Testing and Materials, 1984.

7. ASTM Manual 7. "Manual on Presentation of Data and Control Chart Analysis." West Conshohocken, PA: American Society for Testing and Materials, 1990.

8. ASTM D1141. "Standard Specification for Substitute Ocean Water." West Conshohocken, PA: American Society for Testing and Materials, 1998• 1.

9. ASTM D3694. "Standard Practice for Preparation of Sample Containers and for Preservation of Organic Constituents." West Conshohocken, PA: American Society for Testing and Materials, 1996.

10. ASTM D3856. "Standard Guide for Good Laboratory Practices in Laboratories Engaged in Sampling and Analysis of Water." West Conshohocken, PA: American Society for Testing and Materials, 1995.

11. ASTM D4220. "Standard Practice for Preserving and Transporting Soil Samples." West Conshohocken, PA: American Society for Testing and Materials, 1995.

12. ASTM D4515. "Standard Practice for Estimation of Holding Time for Water Samples Containing Organic Constituents." West Conshohocken, PA: American Society for Testing and Materials, 1985(1995)• 1.

13. ASTM D4840. "Standard Guide for Sampling Chain-of-Custody Procedures." West Conshohocken, PA: American Society for Testing and Materials, 1995.

14. ASTM D5905. "Standard Practice for the Preparation of Substitute Wastewater." West Conshohocken, PA: American Society for Testing and Materials, 1998• 1.

15. ASTM D6051. "Standard Guide for Composite Sampling and Field Subsampling for Environmental Waste Management Activities." West Conshohocken, PA: American Society for Testing and Materials, 1996.

16. ASTM D6362. "Standard Practice for Certificates of Reference Materials for Water Analysis." West Conshohocken, PA: American Society for Testing and Materials, 1998.

17. ASTM E11. "Standard Specification for Wire Cloth and Sieves for Testing Purposes." West Conshohocken, PA: American Society for Testing and Materials, 1995.

18. ASTM E200. "Standard Practice for Preparation, Standardization, and Storage of Standard and Reagent Solutions for Chemical Analysis." West Conshohocken, PA: American Society for Testing and Materials, 1997.

19. ASTM E548. "Standard Guide for General Criteria Used for Evaluating Laboratory Competence." West Conshohocken, PA: American Society for Testing and Materials, 1994• 1.

20. ASTM E691. "Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method." West Conshohocken, PA: American Society for Testing and Materials, 1999.

21. ASTM E1323. "Standard Guide for Evaluating Laboratory Measurement Practices and the Statistical Analysis of the Resulting Data." West Conshohocken, PA: American Society for Testing and Materials, 1989(1994)• 1.

22. ASTM E1601. "Standard Practice for Conducting an Interlaboratory Study to Evaluate the Performance of an Analytical Method." West Conshohocken, PA: American Society for Testing and Materials, 1998.

23. ASTM E1763. "Standard Guide for Interpretation and Use of Results from Interlaboratory Testing of Chemical Analysis Methods." West Conshohocken, PA: American Society for Testing and Materials, 1998.

24. Duncan, Acheson J. "Quality Control and Industrial Statistics." Homewood, IL: Richard D. Irwin, Inc., 1986.

25. Gilbert, Richard O. "Statistical Methods for Environmental Pollution Monitoring." New York: John Wiley & Sons, Inc., 1987.

26. Grant, Eugene I. and Richard S. Leavenworth. "Statistical Quality Control." New York: McGraw-Hill, Inc., 1996.

27. ISO 3534-1. "Statistics - Vocabulary and Symbols - Part 1: Probability and General Statistical Terms." Geneva, Switzerland: International Organization for Standardization, 1993.

28. ISO 3534-2. "Statistics - Vocabulary and Symbols - Part 2: Statistical Quality Control." Geneva, Switzerland: International Organization for Standardization, 1993.

29. ISO 3534-3. "Statistics - Vocabulary and Symbols - Part 3: Design of Experiments." Geneva, Switzerland: International Organization for Standardization, 1993.

30. ISO 8402. "Quality Management and Quality Assurance - Vocabulary." Geneva, Switzerland: International Organization for Standardization, 1994.

31. ISO 9004-1. "Quality Management and Quality System Elements - Part 1: Guidelines." Geneva, Switzerland: International Organization for Standardization, 1994.

32. ISO 9004-2. "Quality Management and Quality System Elements - Part 2: Guidelines for Services." Geneva, Switzerland: International Organization for Standardization, 1991.

33. ISO 9004-3. "Quality Management and Quality System Elements - Part 3: Guidelines for Processed Materials." Geneva, Switzerland: International Organization for Standardization, 1993.

34. ISO 9004-4. "Quality Management and Quality System Elements - Part 4: Guidelines for Quality Improvement." Geneva, Switzerland: International Organization for Standardization, 1993.

35. ISO 10012-1. "Quality Assurance Requirements for Measuring Equipment - Part 1: Metrological Confirmation System for Measurement Equipment." Geneva, Switzerland: International Organization for Standardization, 1992.

36. ISO 10012-2. "Quality Assurance Requirements for Measuring Equipment - Part 2: Guidelines for Control of Measurement Processes." Geneva, Switzerland: International Organization for Standardization, 1997.

37. ISO Guide 32. "Calibration in Analytical Chemistry and Use of Certified Reference Materials." Geneva, Switzerland: International Organization for Standardization, 1997.

38. ISO Guide 33. "Uses of Certified Reference Materials." Geneva, Switzerland: International Organization for Standardization, 1989.

39. ISO Guide 39. "General Requirements for the Acceptance of Inspection Bodies." Geneva, Switzerland: International Organization for Standardization, 1988.

40. ISO/IEC Guide 57. "Guidelines for the Presentation of Inspection Results." Geneva, Switzerland: International Organization for Standardization, 1991.

41. ISO/IEC Guide 58. "Calibration and Testing Laboratory Accreditation Systems - General Requirements for Operation and Recognition." Geneva, Switzerland: International Organization for Standardization, 1993.

42. ISO/IEC Guide 61. "General Requirements for Assessment and Accreditation of Certification/Registration Bodies." Geneva, Switzerland: International Organization for Standardization, 1996.

43. ISO/IEC Guide 62. "General Requirements for Bodies Operating Assessment and Accreditation of Certification/Registration of Quality Systems." Geneva, Switzerland: International Organization for Standardization, 1996.

44. ISO/IEC Guide 66. "General Requirements for Bodies Operating Assessment and Certification/Registration of Environmental Management Systems (EMS)." Geneva, Switzerland: International Organization for Standardization, 1996.

45. ISO/TR Guide 13425. "Guide for the Selection of Statistical Methods in Standardization and Specification." Geneva, Switzerland: International Organization for Standardization, 1995.

46. ISO "Directory of Certified Reference Materials." Geneva, Switzerland: International Organization for Standardization, 1982.

47. ISO Handbook 3. "Statistical Methods." Geneva, Switzerland: International Organization for Standardization, 1981.

48. ISO Handbook. "Statistical Method for Quality Control, Volume 1: Terminology and Symbols; Acceptance Sampling." Geneva, Switzerland: International Organization for Standardization, 1995.

49. ISO Handbook. "Statistical Method for Quality Control, Volume 2: Measurement Methods and Results; Interpretation of Statistical Data; Process Control." Geneva, Switzerland: International Organization for Standardization, 1995.

50. Kafadar K. "A Biweight Approach to the One-Sample Problem." *Journal of the American Statistical Association*, 77, 1982, pp. 416-424.

51. NIOSH Publication 95-104. "Laboratory Evaluations and Performance Reports for the Proficiency Analytical Testing (PAT) and Environmental Lead Proficiency Analytical Testing (ELPAT) Programs." Cincinnati, OH: National Institute for Occupational Safety and Health, November 1994.

52. NIST Handbook 145. "Handbook for the Quality Assurance of Metrological Measurements." Gaithersburg, MD: National Institute of Standards and Technology, 1986.

53. NIST Handbook 150. "National Voluntary Laboratory Accreditation Program, Procedures and General Requirements." Gaithersburg, MD: National Institute of Standards and Technology, 1994.

54. NIST Handbook 150-xx, "National Voluntary Laboratory Accreditation Program, Chemical Calibration, Providers of Proficiency Testing." Gaithersburg, MD: National Institute of Standards and Technology, Draft, October 1998.

55. NIST Internal/Interagency Reports 5589 Rev. "National Voluntary Laboratory Accreditation Program - Quality System Manual." Gaithersburg, MD: National Institute of Standards and Technology, October 1997.

56. Royal Society of Chemistry. "Robust Statistics - How not to Reject Outliers, Part 1, Basic Concepts, and Part 2, Interlaboratory Trials." *Analyst*, 114, 1989, pp. 1693-1702.

57. Taylor, John K. "Quality Assurance of Chemical Measurements." Chelsea, MI: Lewis Publishers, Inc., 1987.

58. Taylor, John K. "Statistical Techniques for Data Analysis." Chelsea, MI: Lewis Publishers, Inc., 1990.

59. Tholen, D. W. "A Statistical Procedure for Measuring and Evaluating Performance in Interlaboratory Comparison Programs." *Arch. Pathol. Lab. Med.*, 112, 1982, pp. 416-424.

60. Tholen, D. W. "Reference Values and Participant Means as Targets in Proficiency Testing." *Arch. Pathol. Lab. Med.*, 117, 1993, pp. 885-889.

61. USACE CRREL Special Report 95-3. "Collaborative Study of Soils Spiked with Volatile Organic Compounds." Washington, DC: United States Army Corps of Engineers, 1995.

62. USACE CRREL Special Report 95-5. "Chemical Preservation of Volatile Organic Compounds in Soil Subsamples." Washington, DC: United States Army Corps of Engineers, 1995.

63. USACE CRREL Special Report 95-26. "Enhanced Preservation of Volatile Organic Compounds in Soil with Sodium Bisulfate." Washington, DC: United States Army Corps of Engineers, 1995.

64. USEPA. "Guide to Method Flexibility and Approval of EPA Water Methods." Washington, DC: United States Environmental Protection Agency, Draft, July 1996.

65. USEPA/530/SW-87/003. "Test Method Equivalency Petitions." Washington, DC: United States Environmental Protection Agency, February 1987.

66. USEPA/600/9-91/034. "Variability in Protocols." Washington, DC: United States Environmental Protection Agency, September 1991.

67. USEPA QA/G-6. "Guidance for the Preparation of Standard Operating Procedures (SOPs) for Quality-Related Documents." Washington, DC: United States Environmental Protection Agency, November 1995.

68. USEPA QA/G-9. "Guidance for Data Quality Assessment: Practical Methods for Data Analysis." Washington, DC: United States Environmental Protection Agency, January 1998.

69. USEPA/SW-846. "Test Methods for Evaluating Solid Waste Physical/Chemical Methods." Washington, DC: United States Environmental Protection Agency, Third Edition, September 1986, including Final Updates I (July 1992), II (September 1994), IIA (August 1993), IIB (January 1995), and III (June 1997).

70. Winer, Benjamin J., Donald R. Brown, and Kenneth M. Michels. "Statistical Principles in Experimental Design." New York: McGraw-Hill, Inc., 1991.

71. Youden, W. J. "Ranking Laboratory by Round-Robin Tests." *Materials Research & Standards*, January 1963.

Appendix B
Abbreviations, Acronyms, and Symbols

AA	Atomic absorption spectroscopy
ACS	American Chemical Society
ANOVA	Analysis of variance
ANSI	American National Standards Institute
AOAC	Association of Official Analytical Chemists
APHA	American Public Health Association
ASQ	American Society for Quality
ASQC	American Society for Quality Control (former ASQ)
ASTM	American Society for Testing and Materials
AWWA	American Water Works Association
BRAC	Base Realignment and Closure
CDQM	Chemical Data Quality Management
CECW-E	Corps of Engineers, Directorate of Civil Works, Engineering and Construction Division
CEMP-RT	Corps of Engineers, Directorate of Military Programs, Environmental Division, Policy and Technology Branch
CFR	Code of Federal Regulations
CHP	Chemical Hygiene Plan
CLP	Contract Laboratory Program
COC	Chain-of-custody
COR	Contracting Officer Representative

CRC	Chemical Rubber Company
CRM	Certified reference materials
CRREL	Cold Regions Research and Engineering Laboratory
CVAA	Cold vapor atomic absorption spectroscopy
DataQUEST	Data Quality Evaluation Statistical Toolbox
DDT	Dichlorodiphenyl trichloroethane
DERP	Defense Environmental Restoration Program
DLA	Defense Logistics Agency
DOD	Department of Defense
DOE	Department of Energy
DOT	Department of Transportation
DQO	Data Quality Objective
ELPAT	Environmental Lead Proficiency Analytical Testing
EM	Engineer Manual
EMS	Environmental Management Systems
Endrin	A trade name for an insecticide
EPA	Environmental Protection Agency
ER	Engineer Regulation
ERDC	Environmental Research and Development Center
FE	Fundamental error
FLAA	Flame atomic absorption spectroscopy

FOA	Field operating activities
FR	Federal Register
FUDS	Formerly Used Defense Sites
GC	Gas chromatography
GC/MS	Gas chromatography/mass spectrometry
GE	Grouping and segregation error
GFAA	Graphite furnace atomic absorption spectroscopy
GHAA	Gaseous hydride atomic absorption spectroscopy
HQUSACE	Headquarters, U.S. Army Corps of Engineers
HTRW	Hazardous, Toxic and Radioactive Waste
HTRW-CX	Hazardous, Toxic and Radioactive Waste Center of Expertise
ICP	Inductively coupled plasma
ICP-AES	Inductively coupled plasma-atomic emission spectroscopy
ICP-MS	Inductively coupled plasma-mass spectrometry
IEC	International Electrotechnical Commission
IRP	Installation Restoration Program
ISO	International Organization for Standardization
MDL	Method Detection Limit
MFR	Memorandum For Record
MQL	Method Quantitation Limit
MRL	Method Reporting Limit

MSC	Major Subordinate Command
MSDC	Mass Spectral Data Center
MSDS	Material Safety Data Sheet
NCSL	National Conference of Standards Laboratories
NELAC	National Environmental Laboratory Accreditation Conference
NIOSH	National Institute for Occupational Safety and Health
NIST	National Institute of Standards and Technology
OEW	Ordnance and Explosive Waste
OSHA	Occupational Safety and Health Administration
OSW	Office of Solid Waste
PAT	Proficiency Analytical Testing
PE	Performance Evaluation
pH	A value representing the acidity or alkalinity of an aqueous solution
PM	Project Manager
PP	Priority Pollutant
QA	Quality Assurance
QC	Quality Control
RCRA	Resource Conservation and Recovery Act
RSD	Relative Standard Deviation
SD	Standard deviation
SE	Standard error

SF	Superfund
SOP	Standard operating procedure
SW-846	USEPA publication, "Test Methods for Evaluating Solid Waste"
TAL	Target Analyte List
TR	Technical report
USACE	U.S. Army Corps of Engineers
USEPA	U.S. Environmental Protection Agency
v/v	Volume to volume ratio
WP	Water Pollution
WS	Water Supply

Appendix C
Soil Cleaning Procedures

C.1. Procedures.

Soils contain a variety of metals and organics that may react with target analytes. Therefore, soil cleaning is usually required. Use mild treatments to preserve the character of the soil. Otherwise, the soil may become sandy and unable to mimic the complexing and absorption nature of real-world samples. The soil would then be an unsuitable matrix for PE samples. Suitable treatment methods described below are acid wash for inorganic contaminants and solvent extraction for organic contaminants.

C.1.1. Acid Wash. Acid wash is usually used to reduce inorganic contaminants such as metals from soils. Follow these steps to acid wash soil:

(1) Weigh 10 grams of soil and place into a 4-liter plastic container along with a Teflon®-coated magnetic stirring bar.
(2) Add 4 liters of pH 2 nitric acid solution and start stirring.
(3) Set the stirrer speed to suspend most of the soil into the acid solution, but not so fast that the stirring bar cannot track the magnet.
(4) Check the pH and add 50% v/v (1+1) nitric acid to bring the pH to less than one if not already at that level.
(5) Check the pH hourly to ensure it does not rise above two. If the pH rises above two, adjust to below pH 2 with 50% v/v (1+1) nitric acid.
(6) Remove the container from magnetic stirrer after stirring for four hours and allow to stand until all particulate matter has settled.
(7) Decant carefully and dispose of the acid solution according to regulations.
(8) Add four liters of deionized water and start the stirring again.
(9) Stop the stirring after one hour and allow the soil to settle to the bottom again.
(10) Decant the wash water and repeat the water washing two or more times.
(11) Check the pH of the final wash water. It should be between pH 4 and pH 6. If the pH is less than 4, additional water washes are required until the pH is within the required range.
(12) Do not add base to adjust the pH under any circumstances.
(13) Filter the washed soil through a 0.45-• m membrane filter and air dry.
(14) Store the treated soil in a nitric acid-washed plastic jar with a Teflon®-lined cap at 4°C in the dark. This treated soil will be used to prepare PE samples for inorganic analyses.

If a large quantity of acid-washed soils is needed, follow these steps:

(1) Treat multiple portions of soils.
(2) Combine all treated portions into one container after the second wash.
(3) Wash the combined soil with deionized water until the pH falls in the required range.
(4) Filter through a 0.45-• m membrane filter and allow to air dry.
(5) Store the treated soil in an acid-washed plastic jar with a Teflon®-lined cap at 4°C in the dark.

Note: The acid wash procedures may be scaled up; however, the proportions of nitric acid solution to soil must remain the same in all cases.

C.1.2. Solvent Extraction. Solvent extraction is usually used to reduce organic contaminants such as semivolatile organic compounds from soils. Because soils do not always contain high concentrations of organic contaminants, it is preferred to collect clean soils from uncontaminated sites rather than to clean dirty soils from contaminated sites. Follow these steps for solvent extraction:

(1) Place 20 grams of dry soils in an extraction thimble or between two plugs of glass wool.
(2) Extract the soils with 300 mL of appropriate solvent (e.g., 1:1 [v/v] acetone/hexane, 10:1 [v/v] toluene/methanol, or methylene chloride) in a Soxhlet extractor for 16 - 24 hours.
(3) Air dry the extracted soils and then thermally treat them at 105 - 125°C overnight to remove the remaining extraction solvents.
(4) Store the treated soil in a clean glass jar with a Teflon®-lined cap at 4°C in the dark. This treated soil will be used to prepare PE samples for organic analyses.
(5) To prepare a large quantity of solvent-extracted soils, individually treated portions may be combined into one container and mixed by tumbling and shaking to provide a uniform blend of treated soils.

Note: The above-mentioned procedures may be scaled up; however, the proportions of extraction solvent to soil should remain the same in all cases.

Appendix D
Determination of Appropriate Particle Size of PE Samples

D.1. Homogeneity and Sampling Errors.

Three ways to reduce sampling errors described below include:

- Sample matrix and method.
- Sample size and consistency.
- Particle size.

D.1.1. Sample Matrix and Method. Homogenization of PE samples is most likely the least controlled and error prone step in an analytical chain. It is especially critical in preparation of individual PE samples subsampled from bulk, solid materials because of direct effects on sample integrity. Homogeneity of PE samples of single-phase liquid matrices is not a problem unless analyte loss to container walls is a concern. Shaking and mixing just prior to analysis are usually sufficient to ensure homogeneity of bulk, aqueous PE sample materials. This appendix therefore focuses on homogeneity of solid PE samples.

D.1.2. Sample Size and Consistency. Solid PE samples should be as finely ground and homogeneous as possible. Smaller, more manageable quantities of PE samples can then be shipped. If laboratory proficiency in subsampling is to be tested, a coarse sample should be shipped. However, sufficient quantities must be furnished to allow execution of proper sampling technique. For example, several kilogram quantities may be necessary for samples with maximum particle size of 0.5 cm. If sufficient quantities are not shipped, the laboratory cannot follow proper sampling technique and achieve analyses that are representative of the lot of PE samples. For obvious reasons, proficiency testing in subsampling can be very expensive, logistically difficult, and is normally not included in the USACE PE Program.

D.1.3. Particle Size. The particle size and the subsample size of solid PE sample materials play important roles in determining sampling errors on contaminant concentrations. Because there is little room to change the method-specified subsample size, the particle size is the variable that can be adjusted to minimize sampling errors. Pierre Gy's Sampling Theory presents a method of estimating sampling precision of particulate materials at specific particle sizes. The theory describes how errors are generated; how they can be eliminated or reduced; and how the residual error can be estimated. The details on Pierre Gy's Sampling Theory are presented in several reference books and are beyond the scope of this appendix. This appendix will focus on "fundamental error" and "grouping and segregation error," which are directly applicable to heterogeneity problems with PE sample materials.

D.2. Pierre Gy's Sampling Theory.

D.2.1. Fundamental Error (FE). According to *Pierre Gy's Sampling Theory and Sampling Practice*, FE results from the constitution heterogeneity of PE sample materials. Constitution heterogeneity, related to individual particles of PE sample materials, is an intrinsic property of the PE sample materials and cannot be varied, unless processed with a comminution. Mixing and homogenization of PE sample materials have no influence on constitution heterogeneity. PE samples prepared by subsampling a bulk PE sample material are affected by an error specifically related to constitution heterogeneity. This error is known as the FE. For PE samples of a given weight, FE is an incompressible minimum depending on intrinsic properties of PE sample materials. FE is the only error that can never be canceled out even with perfect sampling operations. It is also the only error that can be estimated beforehand. FE is related to the sample size and particle size. It can be minimized for each PE sample material through reduction in particle size of PE sample materials.

D.2.2. Grouping and Segregation Error (GE). GE results from the distribution heterogeneity of PE sample materials. When preparing individual PE samples, individual particles making up PE samples are not collected strictly at random or one by one. Increments that are likely made of many particles are collected to make up individual PE samples. Statistically speaking, a PE sample is not made of strictly random particles, but only of random groups of particles. When all groups of particles that may be subsampled have the same average composition, a PE sample material is homogeneous in practice. Otherwise, the PE sample material has a heterogeneous distribution that leads to GE. The distribution heterogeneity is directly proportional to constitution heterogeneity and grouping and segregation factors that are naturally introduced by sampling process and gravitational force. Distribution heterogeneity can be minimized by mixing; however, there is always a residual heterogeneity that never goes to zero and is a characteristic of the PE sample material itself. GE is often negligible with major constituents but could become important with minor constituents. According to Gy, when optimizing sampling protocol, one may always assume that the GE is equal to FE. The overall sampling error is therefore the sum of FE and GE.

D.2.3. Estimating Sampling Errors. Major sample preparation efforts, including reduction in particle size and distribution heterogeneity, are designed to minimize FE and GE. The relationship between FE, sample size, and particle size is as follows.

$$S^2 = \frac{18 \times f \times e \times d^3}{M_s}$$

where:
S^2 = relative variance of contaminant concentration due to FE
S = relative standard deviation of contaminant concentration due to FE
f = a dimensionless factor related to particle shape (a typical value is 0.5)
e = average density in grams per cubic centimeter (ca. 2.5 g/cm^3 for soil)
d = diameter of the largest particle in centimeters
M_s = sample mass in grams

D.2.4. Maximum Allowable Particle Size. Using the above assumptions for density and shape factor, this equation can be rearranged to calculate the largest particle size that can be representatively accommodated by a given subsample mass and FE.

$$d = \sqrt[3]{\frac{M_s \times S^2}{22.5}}$$

If the density of bulk PE sample materials being sampled varies significantly from 2.5 g/cm^3, the actual density should be used. If PE sample material does not have a typical spherical shape, the factors in Table D-1 can be substituted for the f factor of 0.5.

Table D-1. "f" Values for Different Particle Shapes

Particle Shape	f
Cubic	1
Spheres	0.5
Flakes	0.1
Soft solids shaped by mechanical stress	0.2
Needles	>1 to • 10

Table D-2 lists the maximum allowable particle size in centimeters that can be accommodated by a given subsample mass at varying FE. The tighter the required precision, the larger the sample mass and/or the smaller the particle size. This table assumes that the soil particle has a typical spherical shape and serves as an example how the maximum allowable particle size changes as a function of these parameters. For proficiency testing, if limited amounts of PE samples are

shipped, keep the total sampling errors including FE and GE as low as practically feasible but no larger than 15%.

Table D-2. Maximum Allowable Particle Size That Can Be Accommodated by a Given Subsample Mass at Various Percent Relative Standard Deviation (RSD)

Recommended Subsample Mass (g)	U.S. Standard Sieve No.	Nominal Sieve Opening (cm)	Maximum Allowable Particle Size (cm)		
			5% RSD	10% RSD	15% RSD
0.01	70	0.0212	0.010	0.016	0.022
0.02	60	0.0250	0.013	0.021	0.027
0.03	50	0.0300	0.015	0.024	0.031
0.05	45	0.0355	0.018	0.028	0.037
0.1	40	0.0425	0.022	0.035	0.046
0.2	35	0.0500	0.028	0.045	0.059
0.3	30	0.0600	0.032	0.051	0.067
0.5	25	0.0710	0.038	0.061	0.079
1	18	0.100	0.048	0.076	0.100
2	16	0.118	0.061	0.096	0.126
3	14	0.140	0.069	0.110	0.144
5	12	0.170	0.082	0.131	0.171
10	10	0.200	0.104	0.164	0.215
20	8	0.236	0.131	0.207	0.271
30	7	0.280	0.149	0.237	0.311
50	6	0.335	0.177	0.281	0.368
100	5	0.400	0.223	0.354	0.464

Appendix E
Statistical Analysis

E.1. Overview.

This appendix describes the statistical procedures that are generally used by the USACE for certification of PE samples. It will not cover the details of experimental design, the derivation of equations, or the justification for use of particular equations. These topics are beyond the scope of this Appendix and are very well covered by a number of standard textbooks.

Statistical procedures along with examples discussed in this Appendix include:

- Distribution test
- Outlier test
- Homogeneity test
- Stability test
- Reproducibility
- Reference value
- Prediction interval

E.2. General Guidelines.

- Use a statistician to assist with all aspects of experimentation.
- Provide the statistician with background information on the proposed experiment so the experimenter and the statistician may determine the best approach to suit the experimental objectives and constraints.
- Consider using statistical tools such as commercially available spreadsheets (e.g., Excel[1]) and statistical software packages (e.g., DataQUEST[2], Minitab[3], SAS[4], etc.) for data handling and analysis.

E.3. Distribution Test.

E.3.1. Purpose. Many statistical models and tests are only appropriate for data that follow a normal distribution or can be transformed into a normal distribution. This is especially true for chemical data. Testing for normality is crucial in selection of appropriate statistical methods.

---

[1] Excel is a registered trademark of Microsoft Corporation, One Microsoft Way, Redmond, WA 98052.
[2] DataQUEST is a free software developed by USEPA, 401 M Street, SW, Washington, DC 20460.
[3] Minitab is a registered trademark of Minitab Inc., 3081 Enterprise Drive, State College, PA 16801.
[4] SAS is a registered trademark of the SAS Institute Inc., Campus Drive, Cary, NC 27513.

E.3.2. Test Types. The test of data distribution can be performed either qualitatively using a normal probability plot, histogram, or stem-and-leaf plot, or quantitatively using a statistical analysis to confirm or reject the assumptions that accompany a statistical test. For normally distributed data, a normal distribution plot approximately follows a straight line. If data are not normally distributed, there are large deviations from a straight line in the tails or middle of a normal distribution plot. Both histogram and stem-and-leaf plots of a normal distribution show bell-shaped curves. Using a plot to decide if the data are normally distributed involves making a subjective decision which is easy to make for extremely non-normal data. When there is no straightforward decision; however, formal quantitative procedures such as W test or Filliben's statistic are usually necessary to test the assumption of normality. W test and Filliben's statistic compute the correlations between the quantiles of the standard normal distribution and the ordered values of sample data. If the data follow a normal distribution curve, the test statistic will be relatively high. However, both tests are difficult to compute manually due to a large number of summations and multiplications. Use USEPA's DataQUEST software to perform both tests. The rest of this section explains the procedure for a normal distribution plot.

E.3.3. Normal Probability Plot.

E.3.3.1. Procedure. Follow these steps to perform a normal probability plot:

1. Order the data ($X_1, X_2, \ldots, X_n$) from the lowest to the highest.
2. Compute the absolute frequency (AF_i), i.e., the number of times each value occurs, for each data value; and the cumulative frequency (CF_i) and Y_i with the following formulas:

$$CF_i = \sum_{i=1}^{n} AF_i$$

$$Y_i = 100 \times \frac{CF_i}{(n + 1)}$$

3. Plot (Y_i, X_i) pairs on a normal probability paper. If the plot of these pairs approximately forms a straight line, the data are probably normally distributed. Otherwise, the data may not be normally distributed.

E.3.3.2. Example. Consider the following eleven data points: 83.0, 86.9, 86.2, 89.7, 80.7, 83.0, 84.1, 87.8, 88.5, 86.2, and 84.5 mg/kg. Rank the data in order and compute the frequencies as shown in Table E-1.

Table E-1. Example of Normal Probability Plot

i	Individual X_i	Absolute Frequency, AF_i	Cumulative Frequency, CF_i	Y_i
1	80.7	1	1	8.33
2	83.0	2	3	25.00
3	84.1	1	4	33.33
4	84.5	1	5	41.67
5	86.2	2	7	58.33
6	86.9	1	8	66.67
7	87.8	1	9	75.00
8	88.5	1	10	83.33
9	89.7	1	11	91.67

A plot of the (Y_i, X_i) pairs using normal probability paper is shown in Figure E.1. Because these pairs apparently form a straight line, the data are probably normally distributed.

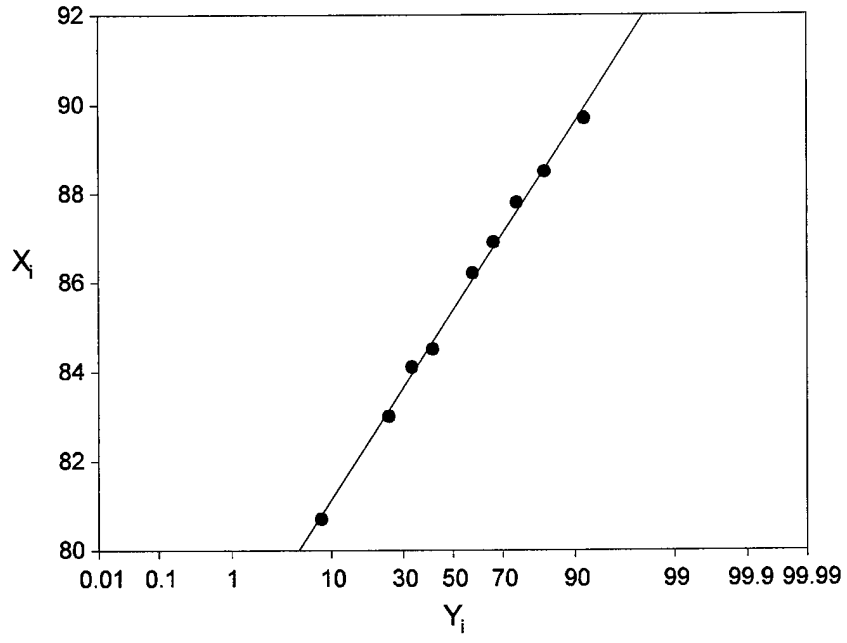

Figure E-1. Normal Probability Plot

E.4. Outlier Test.

E.4.1. Purpose. Outliers are measurements that are extremely large or small compared with the rest of the sample data and are suspected of misrepresenting the population from which they were collected. Statistical outlier tests provide probabilistic evidence that the extreme values do not fit with the distribution of the remainder of the data and are therefore statistical outliers. These tests should only be used to identify data points that require further investigations. The tests alone cannot determine if a statistical outlier should be discarded or corrected. This decision should be based on scientific and judgmental grounds, in addition to the results of statistical outlier tests.

E.4.2. Test Types. USEPA's DataQUEST software package provides several popular tests including Grubbs' tests, Dixon's test, Rosner's test, and Walsh's test. Select a test based on sample size, data distribution, the number of outliers, etc. Follow these guidelines also:

- Use Grubbs' test when data are normally distributed and the sample size is not greater than 50.
- Use Rosner's test if sample size is equal to or greater than 50.
- Use a nonparametric test, such as Walsh's test, if the data are not normally distributed or cannot be transformed.

An example calculation of the commonly used Grubbs' test is presented below.

E.4.3. Grubbs' Test.

E.4.3.1. Procedure. Three similar tests to predict outliers in normally distributed data were developed by Frank Grubbs. The specific Grubbs' test usually used by the USACE PE Program considers the smallest and/or largest value(s) of the data set as the suspected outlier(s) and is discussed here.

1. Perform a normality test of underlying data distribution without the suspected outliers prior to performing a Grubbs' test.
2. If data pass a normality test, rank data from the smallest to the largest to detect any suspected outliers.
3. Compute the sample mean and standard deviation (*SD*) according to the following formulas:

$$mean = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$SD = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n - 1}}$$

$$= \sqrt{\frac{\sum_{i=1}^{n} X_i^2 - \frac{1}{n} \left(\sum_{i=1}^{n} X_i \right)^2}{n - 1}}$$

If the suspected outlier is the smallest value of the data set, the test statistic of the Grubbs' test is:

$$G = \frac{\bar{X} - X_1}{SD}$$

If the suspected outlier is the largest value of the data set, the test statistic is:

$$G = \frac{X_n - \bar{X}}{SD}$$

If G exceeds the critical value in Table F-1 of Appendix F, either X_1 and/or X_n is the outlier depending on the test statistic.

E.4.3.2. Example. Upon examining the ordered sample data: 82, 87, 92, 98, 103, 105, 106, 108, 113, and 151 mg/L, one suspects that the largest value (151 mg/L) of the data set could be an outlier. Because the mean value and standard deviation are sensitive to data distribution, a normality test is first performed on the data set. A normal probability plot shows that there is no reason to suspect that the data without suspected outliers are not normally distributed. Based on the mean value of 104.5 mg/L and standard deviation of 19.04 mg/L:

$$G = \frac{X_n - \bar{X}}{SD} = \frac{151 - 104.5}{19.04} = 2.44$$

Because G = 2.44, which is greater than 2.176 from Table F-1 of Appendix F, there is evidence that the largest value, 151, is an outlier at a 0.05 significance level and should be further investigated.

E.5. Homogeneity Test.

E.5.1. Procedure. To test the homogeneity of bulk PE sample material in a container, follow these steps:

1. Collect five samples randomly from each of the top, middle, and bottom sections of the container to yield a total of 15 samples. The globe null hypothesis is as follows:

$$H_O: \bullet_1 = \bullet_2 = \bullet_3 = \ldots = \bullet_{15}$$

where • 's are the population means that can be represented by the samples. The alternative hypothesis is:

H_A: The • 's are not all equal.

2. Carry out an ANOVA F test at • = 0.05 to test the globe null hypothesis following these steps:

- Calculate the sum and mean using following formulas.

$$sum = \sum_{j=1}^{n_i} y_{ij}$$

$$mean = \bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

where:
y_{ij} = sample j in section i
n_i = number of samples in section i
sum = total amount of analyte in section i
mean = mean amount of analyte in section i

- Calculate the ANOVA quantities of within, between, and total Sum of Squares, *df*'s, and Mean Squares using following formulas.

$$\textit{Sum of Squares (within)} = \sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

$$\textit{Sum of Squares (between)} = \sum_{i=1}^{m} n_i (\bar{y}_i - \bar{y})^2$$

$$\textit{Sum of Squares (total)} = \textit{Sum of Squares (within)} + \textit{Sum of Squares (between)}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

$$\bar{y} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n_i} y_{ij}}{\sum_{i=1}^{m} n_i}$$

$$df \textit{ (within)} = (\sum_{i=1}^{m} n_i) - m$$

$$df \textit{ (between)} = m - 1$$

$$df \textit{ (total)} = df \textit{ (within)} + df \textit{ (between)} = (\sum_{i=1}^{m} n_i) - 1$$

$$\textit{Mean Square (within)} = \frac{\textit{Sum of Squares (within)}}{\textit{df (within)}}$$

$$\textit{Mean Square (between)} = \frac{\textit{Sum of Squares (between)}}{\textit{df (between)}}$$

where:
m = number of sections
df = degrees of freedom
Sum of Squares (within) = within-section variability
Sum of Squares (between) = between-section variability
Sum of Squares (total) = total variability
Mean Square = Sum of Squares divided by *df*

- Calculate the test statistic of F test as follows:

$$F = \frac{\textit{Mean Square (bewteen)}}{\textit{Mean Square (within)}}$$

- Obtain critical values from an F distribution (Table F-2 of Appendix F) with "Numerator *df*" and "Denominator *df*" equal to *df* (between) and *df* (within), respectively.

3. Interpret results: If F is less than the critical value at a significance level of 0.05, the material is considered homogeneous. If F is greater than the critical value but the total standard deviation of samples, i.e., the square root of Sum of Squares (total), is less than $0.3\bullet$, where $\bullet$ is the target standard deviation of the proficiency test, the material may still be regarded as sufficiently homogeneous. Otherwise, the material must be reprocessed and rechecked for homogeneity, or an alternative material selected. The only other approach would be to relax the target standard deviation for that particular material, i.e., to take account of the variance that the material will contribute to the results of individual participant laboratories. However, such a practice would destroy the utility of the proficiency test and the confidence of

participating laboratories and is not encouraged except for minor deviations from sufficient homogeneity.

E.5.2. Example. Five samples are randomly collected from each of the top, middle, and bottom sections of bulk PE sample material. Each sample is independently, randomly analyzed in duplicate and the mean values of the duplicate results are tabulated in Table E-2.

Table E-2. Example of Homogeneity Test

Section	Top	Middle	Bottom
Sample Data	85.5	84.0	79.0
	84.5	87.0	89.5
	83.5	82.0	82.5
	79.5	89.5	87.0
	83.0	86.5	85.0
n_i	5	5	5
sum	416	429	423
mean	83.2	85.8	84.6

Calculate the ANOVA quantities for analytes in the three sections and present them in an ANOVA table. The values in Table E-3 can be calculated with a calculator or statistical software.

Table E-3. ANOVA Table for Homogeneity Test

Source	Sum of Squares	*df*	Mean Square
Between sections	16.95	2	8.48
Within sections	119.80	12	9.98
Total	136.75	14	

Calculate the F test statistic as follows:

$$F = \frac{\textit{Mean Square (between)}}{\textit{Mean Square (within)}} = \frac{8.48}{9.98} = 0.850$$

Because F = 0.850 is less than the critical value 3.89 at a significance level of 0.05 with *df* (between) = 2 and *df* (within) = 12, H_0 is not rejected. There is insufficient evidence to conclude any heterogeneity among different sections of PE sample material with respect to the mean concentration. The observed differences in mean concentrations can readily be attributed to chance variation.

E.6. Stability Test.

E.6.1. Test Types. Two approaches will be used to test the stability of PE samples. For PE samples of low stability and short holding times, The following test procedure applies:

1. Conduct duplicate analysis of at least five samples, randomly selected from the production run at the beginning and end of the proficiency test period.
2. Check the equality of variances of two population means with an appropriate statistical test, such as F test, Bartlett's test, Levene's test, etc.
3. If the variances of the two populations are approximately equal, compare the results at the end of the test period with the results at the beginning of the test period using a conventional two-sample *t* test. The mean values shall not be statistically different at • = 0.05 level.
4. If the two populations have unequal variances, compare the means with other tests such as Satterthwaite's two-sample *t* test.

For PE samples of high stability and longer holding time, use a trend test. See Section E.6.5 for procedures and description. The remainder of this section describes F Test, two-sample *t* test, and Satterthwaite's two-sample *t* test.

E.6.2. F Test.

E.6.2.1. Purpose and Underlying Assumptions. Use an F test to determine if the underlying variances of two populations are equal prior to a two-sample *t* test for equality of means. The assumptions underlying an F test are that the two samples are independent, random samples of normal distributions. Confirm these assumptions in the following ways:

- To determine if the samples are independent and random, review the sampling procedures.
- To determine normality, consider sample size. If both sample sizes are large, assume normality without further verifications. For small sample sizes, test the normality of each sample following the procedures in Section E.3.
- Nevertheless, a two-sample *t* test is robust to deviation from the normality of samples and equality of variances. In addition, because the stability test is usually performed by one laboratory, the variances of the two populations are usually equal.

E.6.2.2. Procedure. To perform an F test, following these steps:

1. Calculate the sample variances, SD_1^2 and SD_2^2, and the test statistic, F ratio, where SD_A^2 is the larger of the two variances.

$$F = \frac{SD_A^2}{SD_B^2}$$

2. Compare F with the critical $F_{1..\alpha/2}$ value at numerator $df = (n_A \bullet 1)$ and denominator $df = (n_B \bullet 1)$ in Table F-2 of Appendix F.
3. Interpret results: If $F < F_{1..\alpha/2}$, there is no evidence that the two variances of the two populations are different.

E.6.2.3. Example. Consider the data from a stability test below. Assume that each sample is independently and randomly collected and analyzed in duplicate. The mean values of the duplicate results are listed in Table E-4.

Table E-4. Example of Stability Test

Sample	Control (X_1) (Time t_0)	Test (X_2) (Time t_x)
	84.0	85.5
	87.0	84.5
	82.0	83.5
	89.5	79.5
	86.5	83.0
n_i	5	5
mean	85.8	83.2
SD_i	2.89	2.28
SD_i^2	8.35	5.20

where:

n_i = sample size (i.e., the number of X_i's)
mean = mean value of samples
SD_i = standard deviation of samples
SD_i^2 = variance of samples

$$F = \frac{SD_A^2}{SD_B^2} = \frac{SD_1^2}{SD_2^2} = \frac{8.35}{5.20} = 1.61$$

From Table F-2 of Appendix F, one determines the critical value, $F_{0.975}$, is 9.60 for numerator $df = 4$ and denominator $df = 4$ at a 0.05 significance level. Since F is less than $F_{0.975}$, there is no evidence that these two variances are different at 95% confidence level.

E.6.3. Student's Two-Sample *t* Test.

E.6.3.1. Purpose and Underlying Assumptions. Student's two-sample *t* test is used to compare the means of two populations when the variances of the two populations are equal. The basic assumptions required for the two-sample *t* test are independent and random sampling and normally distributed data. The two-sample *t* test is robust to violations of the assumptions of normality and equality of variances. However, if the assumptions of normality and equality of variances have been tested and rejected, use nonparametric methods such as Wilcoxon Rank Sum Test. Because sample means and standard deviations are used in the test, a two-sample *t* test is not robust to outliers.

E.6.3.2. Procedure. Follow these steps to complete a two-sample *t* test:

1. Use at least five samples, randomly selected from the production run at the beginning and end of a proficiency testing study. The null hypothesis is: H_O: $\bullet_1 \bullet \bullet_2$, where $\bullet_1$ and $\bullet_2$ are the mean analyte concentrations at time 1 (Time t_0) and 2 (Time t_x). The alternative hypothesis is: H_A: $\bullet_1 \bullet \bullet_2$. A conventional two-sample *t* test of the null hypothesis is normally carried out with $\bullet = 0.05$.
2. Analyze each sample in duplicate randomly and independently.
3. Calculate the means and standard deviations of the duplicate results. See the following example in Table E-5.

Table E-5. Example of Student's Two-Sample *t* Test

Sample	Control (X_1) (Time t_0)	Test (X_2) (Time t_x)
	84.0	85.5
	87.0	84.5
	82.0	83.5
	89.5	79.5
	86.5	83.0
n_i	5	5
mean	85.8	83.2
SD_i	2.89	2.28

4. Be sure the variances of the two populations are equal. Because the stability testing is usually performed by one laboratory, it would be expected that sample data are of equal or nearly equal n_i and variances. The equality of variances can be checked with an F test as addressed in Section E.6.2. If the variances of the two samples are not equal, use Satterthwaite's t test (see Section E.6.4).
5. Compute the pooled standard deviation (SD_P) as follows. (SD_P below uses SD_i values from the example table.)

$$SD_P = \sqrt{\frac{(n_1 - 1)\, SD_1^2 + (n_2 - 1)\, SD_2^2}{n_1 + n_2 - 2}} = 2.60$$

where: n_i = sample size (i.e., the number of X_i's)
SD_i = standard deviation of samples

6. Compute the t test value and degrees of freedom using these formulas. (Computed t value uses SD_P calculated previously. Degrees of freedom is calculated using data from the example.)

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{SD_P \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = 1.58$$

$$df = (n_1 + n_2 - 2) = 8$$

7. Use Table F-3 of Appendix F to find that the $t_{0.025}$ (i.e., two-tailed 5% critical value) with $df = 8$ is 2.306. Because $t \bullet t_{0.025}$, there is not enough evidence to reject the null hypothesis, i.e., the data do not provide sufficient evidence to claim that the PE samples are unstable.

E.6.4. Satterthwaite's *t* Test.

E.6.4.1. Purpose and Underlying Assumptions. Use Satterthwaite's t test to compare the means of two populations when the variances of the two populations are not equal. The test is demonstrated below with the same example as the two-sample t test (assuming unequal variances).

1. Use at least five samples, randomly selected from the production run at the beginning and end of the proficiency testing study. The same null hypothesis applies: H_O: $\bullet_1 \bullet \bullet_2$, where $\bullet_1$ and $\bullet_2$ are the mean analyte concentrations at time 1 (Time t_0) and 2 (Time t_x). The alternative hypothesis is: H_A: $\bullet_1 \bullet \bullet_2$. Satterthwaite's t test with $\bullet = 0.05$ is carried out to test the null hypothesis.
2. Analyze each sample in duplicate independently and randomly.
3. Calculate the mean values of the duplicate results.

See the following example of Satterthwaite's t test in Table E-6.

Table E-6. Example of Satterthwaite's t Test

Sample	Control (X_1) (Time t_0)	Test (X_2) (Time t_x)
	84.0	85.5
	87.0	84.5
	82.0	83.5
	89.5	79.5
	86.5	83.0
n_i	5	5
mean	85.8	83.2
SD_i	2.89	2.28
SE_i	1.29	1.02

where: n_i = sample size (i.e., the number of X_i's)
mean = mean value of samples
SD_i = standard deviation of samples
SE_i = standard error of mean

The standard error of the mean, SE_i, is defined as:

$$SE_i = \frac{SD_i}{\sqrt{n_i}}$$

4. Calculate the standard error of the difference between two sample means using the formula below.

$$SE_{(\bar{X}_1 - \bar{X}_2)} = \sqrt{SE_1^2 + SE_2^2} = \sqrt{\frac{SD_1^2}{n_1} + \frac{SD_2^2}{n_2}} = 1.65$$

5. Compute the test statistic for Satterthwaite's t test using the formula below. (Computed t value uses SE calculated previously. Degrees of freedom is calculated using data from the example.)

$$t = \frac{(\bar{X}_1 - \bar{X}_2)}{SE_{(\bar{X}_1 - \bar{X}_2)}} = \frac{(85.8 - 83.2)}{1.65} = 1.58$$

$$df = \frac{(SE_1^2 + SE_2^2)^2}{\frac{SE_1^4}{(n_1 - 1)} + \frac{SE_2^4}{(n_2 - 1)}} = 7.56$$

6. Determine the $t_{0.025}$ (i.e., two-tailed 5% critical value) of Student's t distribution with $df = 7.56$ by computer software or estimate it from Table F-3 of Appendix F, with df given by the smaller of $(n_1 \bullet 1)$ and $(n_2 \bullet 1)$, or by $(n_1 + n_2 \bullet 2)$. The former estimate is somewhat conservative (i.e., the true confidence level is slightly greater than 95% when $t_{0.025}$ is used), and the later estimate is somewhat liberal; however, both usually have little effect on the final decision. If $df = (n_1 + n_2 \bullet 2) = 8$, the critical value, $t_{0.025}$, is 2.306, which is greater than the observed t value of 1.58. Because the significance level of the data is greater than $\bullet = 0.05$, the data are judged compatible with H_O; H_O is not rejected, and the data do not provide sufficient evidence to claim that the PE samples are unstable.

E.6.5. Trend Test. Specific procedures, calculation method, and an example of statistical trend tests follows in this section.

E.6.5.1. Procedure. For PE samples of high stability and longer holding time such as real-world PE samples, use an alternative approach to check the stability of PE sample material. Follow these guidelines:

1. Analyze in duplicate at least five samples, randomly selected from different sections of the bulk material.
2. Perform the analysis in the beginning and at the end of the proficiency test period.
3. Compare the results at the end of the test period with the results from the beginning of the test period. The mean shall not be statistically different at • = 0.05 level, using a conventional *t* test as previously described.
4. Test the bulk material at a regular time interval over an extended time period to cover the anticipated shelf-life or until the supply of the PE sample material is exhausted. (Base the anticipated shelf-life on prior experience and/or evidence from technical literature.)

5. Control chart test data and use to determine and monitor the stability, according to ISO 7870, "Control Charts - General Guide and Introduction;" ISO 7966, "Acceptance Control Charts;" ISO 8258, "Shewhart Control Charts;" and ASTM, "Manual on Presentation of Data and Control Chart Analysis," 1990.

E.6.5.2. Statistical Trend Tests. These tests, such as the Mann-Kendall trend test, are very helpful in detection of any trends in measured concentrations over time and in prevention of potential problems with sample contamination or degradation. The basic Mann-Kendall trend test involves listing the sample data in a temporal order, and computing all differences that may be formed between each datum and its earlier data across a triangular table. If there is an underlying upward or downward trend, then these differences will tend to be sufficiently large positive or negative values to suggest the presence of an upward or downward trend, respectively. Zero differences are not included in the test statistic and therefore should be avoided, if possible, by recording data to sufficient accuracy. For censored data of nondetects, a value of half of MDL value shall be assigned to data reported as below the MDL. The basic Mann-Kendall trend test could be applied to small sample sizes (i.e., fewer than ten). For larger sample sizes, a normal approximation to Mann-Kendall test can be used and is described in a number of textbooks. The procedure and an example of the basic Mann-Kendall trend test are presented below.

E.6.5.3. Mann-Kendall Calculation. The Mann-Kendall statistic, S, is calculated as follows:

1. Calculate the total number of positive signs minus the total number of negative signs across in the triangular table.
2. Determine probability "p" using Table F-4 of Appendix F, sample size n, and absolute value of Mann-Kendall statistic S.

3. Interpret results: A positive S indicates a potential upward trend and a negative S indicates a potential downward trend. For an upward trend, reject H_0 (i.e., no trend) if $S > 0$ and $p < \bullet$. For a downward trend, reject H_0 if $S < 0$ and $p < \bullet$.

E.6.5.4. Example. Consider the following five stability test data points listed in a chronological order: 80, 64, 88, 48, and 50 mg/L. The null hypothesis, H_0: No trend, versus the alternative hypothesis H_A: Either an upward or downward trend at $\bullet = 0.05$ significance level is used for a trend test of sample stability. After examining the data listed below in Table E-7, the suspected trend is obviously a downward trend. The triangular table below is constructed to list all possible differences. The sum of signs of the differences across the rows is shown in the last two columns.

Table E-7. Example of Trend Test

Time	1	2	3	4	5	No. of	No. of
Data	80	64	88	48	50	"+" Signs	"•" Signs
80		•	+	•	•	1	3
64			+	•	•	1	2
88				•	•	0	2
48					+	1	0
						3	7

The Mann-Kendall statistic S = (number of "+" signs) • (number of "• " signs) = 3 • 7 = • 4. Based on Table F-4 of Appendix F, the probability, p, is 0.242 with $n = 5$, and •S• = 4. Because S = • 4 < 0 and p = 0.242 • • (0.05), the null hypothesis is not rejected and there is not sufficient evidence to conclude that there is a downward trend in concentration or instability of the PE sample at a 0.05 significance level.

E.7. Reproducibility.

E.7.1. Procedure. A within-batch reproducibility test is similar to a homogeneity test. It consists of the following steps:

1. Run independent tests of multiple PE samples randomly selected from a production batch. Test each sample in duplicate or triplicate, yielding a minimum of 15 tests.
2. Use a traditional ANOVA F test to test the significance of within-batch component of variances as for a homogeneity test.
3. Interpret results: If the F test is not significant at $\bullet = 0.05$ level, the PE samples may be considered equivalent and the production process reproducible. If the F test is significant, conduct an additional test to assure that the significance represents a difference that truly could affect the evaluation of proficiency testing results. To do this, compare the size of

within-batch component of variances with the acceptance limits of PE samples. If the within-batch component is less than 10% of the acceptance limits (i.e., 3•), the within-batch PE samples are considered sufficiently reproducible.

E.8. Reference Value.

E.8.1. Procedure. A number of methods are available to PE Sample Suppliers to establish the reference value for the analyte. They are similar to the procedures required to assign a concentration value to a reference material and are described below:

1. Set the reference value at the mean value of interlaboratory study if the mean value is within the method-specified control limits on bias. The value should be compatible with the prepared value and/or the mean of referee laboratories with a conventional *t* test at • = 0.05 significance level, depending on the specific type of analytes and matrices. (See Section 5.4 of Chapter 5 for procedures to determine reference values of PE samples.)
2. Determine if the underlying population distribution is single modal with a normal distribution prior to computing mean value. Use graphic presentations such as a histogram plot, frequency plot, stem-and-leaf plot, ranked data plot, quantile plot, normal probability plot, etc. to determine this.
3. Perform an outlier test to determine if an extreme value is a statistical outlier and if the statistical outlier should be excluded or modified prior to calculation of mean value. (A mean value is sensitive to extreme values and nondetects.)
4. Set concentrations for nondetects at half of the sample-specific detection limits if the limits meet the study requirements.
5. Perform an outlier test to decide if the estimated value for a nondetect should be included for computation of sample mean.

E.8.3. Median. Sample median (i.e., 50% point) has recently become very popular in replacing sample mean when robust statistics is used. Sample median is not sensitive to extreme values and can easily be used to handle censored data, i.e., nondetect. The USACE PE Sample Program is not currently using robust statistics, but may consider it for specific PE samples on a case-by-case basis. The discussion of robust statistics is beyond the scope of this Appendix.

E.9. Prediction Intervals.

E.9.1. Definition. A sample mean is an estimate of the unknown population mean, • , but differs from it because of sampling fluctuations. However, it is possible to construct a statistical interval known as a confidence interval to contain the population mean with a specific probability which is known as the associated confidence level. Thus a 95% confidence interval on the population mean is an interval which contains • with a probability of 0.95, or, over a large number of samples, the 95% confidence interval will contain the unknown population mean 95%

of the time. Most textbooks on statistics devote extensive space to confidence intervals on population parameters. A prediction interval estimates the range of variation of the observations in a future sample and is used to obtain limits to contain all of a small number of future data based on previous n data and a specified probability. In proficiency testing, it is the interval within which the next laboratory performance is expected to be located based on the results of n prior laboratory proficiency tests. For the USACE PE Program, the prediction intervals that will contain the next proficiency test result at 95% and 99% confidence intervals are used to establish the acceptance limits for proficiency testing.

E.9.2. Calculation. Assume that n independent and random data are available. The two-sided prediction interval can be constructed from the sample mean and standard deviation (SD) as follows.

$$\bar{X} \pm k(n,\ 1-\alpha) \times SD$$

where: $k(n,\ 1-\alpha)$ in Table F-5 of Appendix F is a factor for determining two-sided $100\times(1-\alpha)\%$ prediction intervals for a single future observation given a previous sample of size n.

For values not tabulated in Table F-5, a conservative approximation for $k(n,\ 1-\alpha)$ is:

$$k(n,\ 1-\alpha) = \sqrt{1 + \frac{1}{n}} \times t(n-1,\ 1-\alpha/2)$$

where: n = size of previous samples
$t(n-1,\ 1-\alpha/2) = [100\times(1-\alpha/2)]$'th percentile of the Students' t distribution with $(n-1)$ degree of freedom

This approximation works satisfactorily for most practical purposes, except for combinations of small n and large α. If the n previous data and the single future datum are all randomly selected from the same normal distribution, one can state with $(1-\alpha)\%$ confidence that the single future datum will be within the calculated prediction interval. A prediction interval is sensitive to the assumption of normality of the data distribution. Other procedures for constructing prediction intervals that make no assumptions about the distribution type are also available; however, they are beyond the scope of this Appendix.

E.9.3. Examples. Consider the following data obtained on a normally distributed parameter based on a random sample from an interlaboratory study: 50.9, 45.8, 49.1, 46.0, and 50.4 • g/L. The sample mean is 48.44 • g/L, and the standard deviation is 2.41 • g/L. The 95% and 99% prediction intervals are:

95%: $48.44 \pm (3.041)(2.41) = 48.44 \pm 7.33$

99%: $48.44 \pm (5.043)(2.41) = 48.44 \pm 12.15$

Thus the next datum from the sample population should fall in an interval of 41.11 • g/L to 55.77 • g/L with a probability of 0.95, and in the interval of 36.29 • g/L to 60.59 • g/L with a probability of 0.99.

Appendix F
Statistical Tables

Table F-1. Critical Values for Grubbs' Test

	Significance Level, •			Significance Level, •	
n	0.01	0.05	*n*	0.01	0.05
			31	3.119	2.759
			32	3.135	2.773
3	1.155	1.153	33	3.150	2.786
4	1.492	1.463	34	3.164	2.799
5	1.749	1.672	35	3.178	2.811
6	1.944	1.822	36	3.191	2.823
7	2.097	1.938	37	3.204	2.835
8	2.221	2.032	38	3.216	2.846
9	2.323	2.110	39	3.228	2.857
10	2.410	2.176	40	3.240	2.866
11	2.485	2.234	41	3.251	2.877
12	2.550	2.285	42	3.261	2.887
13	2.607	2.331	43	3.271	2.896
14	2.659	2.371	44	3.282	2.905
15	2.705	2.409	45	3.292	2.914
16	2.747	2.443	46	3.302	2.923
17	2.785	2.475	47	3.310	2.931
18	2.821	2.504	48	3.319	2.940
19	2.854	2.532	49	3.329	2.948
20	2.884	2.557	50	3.336	2.956
21	2.912	2.580			
22	2.939	2.603			
23	2.963	2.624			
24	2.987	2.644			
25	3.009	2.663			
26	3.029	2.681			
27	3.049	2.698			
28	3.068	2.714			
29	3.085	2.730			
30	3.103	2.745			

Table F-2. Percentiles of the F Distribution

Denominator *df*	1••	Numerator *df* 1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	•
1	0.90	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2	61.2	61.7	62.3	62.8	63.1	63.3
	0.95	161	200	216	225	230	234	237	239	241	242	246	248	250	252	253	254
	0.975	648	800	864	900	922	937	948	957	963	969	985	993	1001	1010	1014	1018
	0.99	4052	5000	5403	5625	5764	5859	5928	5981	6022	6056	6157	6209	6261	6313	6339	6366
2	0.90	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.42	9.44	9.46	9.47	9.48	9.49
	0.95	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5
	0.975	38.5	39.0	39.2	39.2	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5
	0.99	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5
3	0.90	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.20	5.18	5.17	5.15	5.14	5.13
	0.95	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.70	8.66	8.62	8.57	8.55	8.53
	0.975	17.4	16.0	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.2	14.1	14.0	13.9	13.9
	0.99	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	26.9	26.7	26.5	26.3	26.2	26.1
4	0.90	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.87	3.84	3.82	3.79	3.78	3.76
	0.95	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.86	5.80	5.75	5.69	5.66	5.63
	0.975	12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.66	8.56	8.46	8.36	8.31	8.26
	0.99	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.2	14.0	13.8	13.7	13.6	13.5
5	0.90	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.39	3.24	3.21	3.17	3.14	3.12	3.11
	0.95	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.62	4.56	4.50	4.43	4.40	4.37
	0.975	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.43	6.33	6.23	6.12	6.07	6.02
	0.99	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.72	9.55	9.38	9.20	9.11	9.02
6	0.90	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.87	2.84	2.80	2.76	2.74	2.72
	0.95	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	3.94	3.87	3.81	3.74	3.70	3.67
	0.975	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.27	5.17	5.07	4.96	4.90	4.85
	0.99	22.8	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.56	7.40	7.23	7.06	6.97	6.88

Table F-2. (continued) Percentiles of the F Distribution

Denominator		Numerator *df*															
df	1••	1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	•
7	0.90	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.63	2.59	2.56	2.51	2.49	2.47
	0.95	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.51	3.44	3.38	3.30	3.27	3.23
	0.975	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.57	4.47	4.36	4.25	4.20	4.14
	0.99	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.31	6.16	5.99	5.82	5.74	5.65
8	0.90	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.46	2.42	2.38	2.34	2.32	2.29
	0.95	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.22	3.15	3.08	3.01	2.97	2.93
	0.975	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.10	4.00	3.89	3.78	3.73	3.67
	0.99	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.52	5.36	5.20	5.03	4.95	4.86
9	0.90	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.34	2.30	2.25	2.21	2.18	2.16
	0.95	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.01	2.94	2.86	2.79	2.75	2.71
	0.975	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.77	3.67	3.56	3.45	3.39	3.33
	0.99	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	4.96	4.81	4.65	4.48	4.40	4.31
10	0.90	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.24	2.20	2.16	2.11	2.08	2.06
	0.95	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.84	2.77	2.70	2.62	2.58	2.54
	0.975	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.52	3.42	3.31	3.20	3.14	3.08
	0.99	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.56	4.41	4.25	4.08	4.00	3.91
12	0.90	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.10	2.06	2.01	1.96	1.93	1.90
	0.95	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.62	2.54	2.47	2.38	2.34	2.30
	0.975	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.18	3.07	2.96	2.85	2.79	2.72
	0.99	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.01	3.86	3.70	3.54	3.45	3.36
15	0.90	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	1.97	1.92	1.87	1.82	1.79	1.76
	0.95	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.40	2.33	2.25	2.16	2.11	2.07
	0.975	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.86	2.76	2.64	2.52	2.46	2.40
	0.99	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.52	3.37	3.21	3.05	2.96	2.87

Table F-2. (continued) Percentiles of the F Distribution

Denominator *df*	1••	Numerator *df* 1	2	3	4	5	6	7	8	9	10	15	20	30	60	120	•
20	0.90	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.84	1.79	1.74	1.68	1.64	1.61
	0.95	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.20	2.12	2.04	1.95	1.90	1.84
	0.975	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.57	2.46	2.35	2.22	2.16	2.09
	0.99	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.09	2.94	2.78	2.61	2.52	2.42
24	0.90	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.78	1.73	1.67	1.61	1.57	1.53
	0.95	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.11	2.03	1.94	1.84	1.79	1.73
	0.975	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.44	2.33	2.21	2.08	2.01	1.94
	0.99	7.82	6.66	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	2.89	2.74	2.58	2.40	2.31	2.21
30	0.90	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.72	1.67	1.61	1.54	1.50	1.46
	0.95	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.01	1.93	1.84	1.74	1.68	1.62
	0.975	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.31	2.20	2.07	1.94	1.87	1.79
	0.99	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.70	2.55	2.39	2.21	2.11	2.01
60	0.90	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.60	1.54	1.48	1.40	1.35	1.29
	0.95	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.84	1.75	1.65	1.53	1.47	1.39
	0.975	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.06	1.94	1.82	1.67	1.58	1.48
	0.99	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.35	2.20	2.03	1.84	1.73	1.60
120	0.90	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.55	1.48	1.41	1.32	1.26	1.19
	0.95	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.75	1.66	1.55	1.43	1.35	1.25
	0.975	5.15	3.80	3.23	2.89	2.67	2.52	2.39	2.30	2.22	2.16	1.95	1.82	1.69	1.53	1.43	1.31
	0.99	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.19	2.03	1.86	1.66	1.53	1.38
•	0.90	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.49	1.42	1.34	1.24	1.17	1.00
	0.95	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.67	1.57	1.46	1.32	1.22	1.00
	0.975	5.02	3.69	3.12	2.79	2.57	2.41	2.29	2.19	2.11	2.05	1.83	1.71	1.57	1.39	1.27	1.00
	0.99	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.04	1.88	1.70	1.47	1.32	1.00

Table F-3. Critical Values of Student's *t* Distribution

					1 • •				
df	0.70	0.75	0.80	0.85	0.90	0.95	0.975	0.99	0.995
1	0.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657
2	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925
3	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.451	5.841
4	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.536	0.691	0.866	1.074	1.340	1.753	2.131	2.602	2.947
16	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
60	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660
120	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617
•	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

Table F-4. Probabilities for the Small-Sample Mann-Kendall Test for Trend

	n					*n*		
S	4	5	8	9	S	6	7	10
0	0.625	0.592	0.548	0.540	1	0.500	0.500	0.500
2	0.375	0.408	0.452	0.460	3	0.360	0.386	0.431
4	0.167	0.242	0.360	0.381	5	0.235	0.281	0.364
6	0.042	0.117	0.274	0.306	7	0.136	0.191	0.300
8		0.042	0.199	0.238	9	0.068	0.199	0.242
10		0.0083	0.138	0.179	11	0.028	0.068	0.190
12			0.089	0.130	13	0.0083	0.035	0.146
14			0.054	0.090	15	0.0014	0.015	0.108
16			0.031	0.060	17		0.0054	0.078
18			0.016	0.038	19		0.0014	0.054
20			0.0071	0.022	21		0.00020	0.036
22			0.0028	0.012	23			0.023
24			0.00087	0.0063	25			0.014
26			0.00019	0.0029	27			0.0083
28			0.000025	0.0012	29			0.0046
30				0.00043	31			0.0023
32				0.00012	33			0.0011
34				0.000025	35			0.00047
36				0.0000028	37			0.00018
					39			0.0000458
					41			0.0000415
					43			0.0000028
					45			0.00000028

Table F-5. Factors for Determining Two-Sided Prediction Intervals for the Next Observation Given Previous Sample of Size n

Previous Sample Size, n	Probability Level		
	90%	95%	99%
2	7.733	15.562	77.964
3	3.372	4.969	11.460
4	2.631	3.558	6.530
5	2.335	3.041	5.043
6	2.176	2.777	4.355
7	2.077	2.616	3.963
8	2.010	2.508	3.711
9	1.961	2.431	3.536
10	1.922	2.372	3.409
11	1.893	2.327	3.310
12	1.869	2.291	3.233
13	1.849	2.261	3.170
14	1.833	2.236	3.118
15	1.819	2.215	3.075
16	1.807	2.197	3.038
17	1.797	2.181	3.006
18	1.788	2.168	2.977
19	1.779	2.156	2.953
20	1.772	2.145	2.932
21	1.766	2.135	2.912
22	1.760	2.127	2.895
23	1.754	2.119	2.880
24	1.749	2.112	2.865
25	1.745	2.105	2.852
26	1.741	2.099	2.840
27	1.737	2.094	2.830
28	1.733	2.088	2.820
29	1.730	2.083	2.810
30	1.727	2.079	2.802

Table F-5. (continued) Factors for Determining Two-Sided Prediction Intervals for the Next Observation Given Previous Sample of Size *n*

Previous Sample Size, *n*	Probability Level		
	90%	95%	99%
31	1.724	2.075	2.794
41	1.704	2.046	2.737
51	1.692	2.029	2.704
61	1.685	2.016	2.682
81	1.674	2.002	2.655
101	1.668	1.994	2.639
201	1.657	1.977	2.607
501	1.650	1.967	2.589
•	1.645	1.960	2.576